Some notes on tensor products

[under construction]

Notes by David Groisser

version date: 10/9/2023

Contents

In these notes, we are interested in tensor products of vector spaces. Unless otherwise specified, "vector space" will mean *real* vector space, for the sake of concreteness. However, (almost?) every definition, construction, etc., that we discuss works equally well for vector spaces over any field.

The *trivial* vector space is the zero-dimensional vector space $\{0\}$; all other vector spaces are called *nontrivial*.

Occasionally the zero-element of a vector space V will denoted 0_V rather than just 0, but the reader is generally expected to realize from context what the notation "0" means, even if it is used several times with different meanings in the same expression, equation, etc.

The term "subspace" is used always in the sense of linear algebra, not topology.

For any vector space V, any $v \in V$, and any $c \in \mathbf{R}$, we allow ourselves to write cv as vc.

The symbol "▲" is used in these notes to mark the end of various non-proof items (e.g. definitions and remarks) when it might be unclear whether the item continues into the next paragraph.

1 The free vector space generated by a set

Let S be any nonempty set.

The set $\mathrm{Func}(S, \mathbf{R}) := \{\text{all functions from } S \text{ to } \mathbf{R}\}$ has a natural vector-space structure in which the zero element is the identically-zero function, and vector-space operations are defined pointwise: for $f, g : S \to \mathbf{R}$ and $c \in \mathbf{R}$, we define $f + g : S \to \mathbf{R}$ and $cf : S \to \mathbf{R}$ by $(f + g)(p) = f(p) + g(p)$ and $(cf)(p) = c\,f(p)$. Henceforth, the notation "$\mathrm{Func}(S, \mathbf{R})$" will denote the corresponding vector space (i.e. the set $\mathrm{Func}(S, \mathbf{R})$, endowed with this canonical vector-space structure), rather than just the underlying *set* of functions.

A function $f : S \to \mathbf{R}$ is said to have *finite support* if $\mathrm{supp}(f) := \{p \in S \mid f(p) \neq 0\}$ is a finite set.[1] Let $\mathbf{R}[S] \subset \mathrm{Func}(S, \mathbf{R})$ denote the set of functions of finite support. It is easily seen that $\mathbf{R}[S]$ is a subspace of $\mathrm{Func}(S, \mathbf{R})$.

For each $p \in S$, define $e_p \in \mathbf{R}[S]$ by

$$e_p(q) := \delta_{p,q} := \begin{cases} 1 & \text{if } q = p \\ 0 & \text{if } q \neq p. \end{cases} \tag{1.1}$$

Then the collection

$$\{e_p\}_{p \in S} \tag{1.2}$$

is a basis of $\mathbf{R}[S]$. For this reason, we call $\mathbf{R}[S]$ the *free vector space generated by*

[1] This coincides with the topological notion of "finite support" if S is given the discrete topology.

S.[2] Observe that the *dimension* of $\mathbf{R}[S]$ is the *cardinality* of S (which might be uncountable).

Some equivalent definitions of $\mathbf{R}[S]$ are:

$$\begin{aligned}\mathbf{R}[S] &= \mathrm{span}(\{e_p\}_{p\in S}) \subset \mathrm{Func}(S,\mathbf{R}) \\ &= \text{subspace of Func}(S,\mathbf{R}) \text{ generated by } \{e_p\}_{p\in S} \\ &= \text{the set of all linear combinations}^3 \text{ of elements of } \{e_p\}_{p\in S} \\ &= \text{smallest subspace of Func}(S,\mathbf{R}) \text{ containing } \{e_p\}_{p\in S}.\end{aligned}$$

Observe that, for every $f \in \mathbf{R}[S]$, we have $f = \sum_{p\in \mathrm{supp}(f)} f(p)\, e_p$. Given any collection of real numbers $\{a_p\}_{p\in S}$ such that $a_p = 0$ for all but finitely many p, let us adopt the convention that $\sum_{p\in S} a_p\, e_p$ means the finite sum $\sum_{\{p\in S\,|\,a_p\neq 0\}} a_p\, e_p \in \mathbf{R}[S]$ (where an empty sum is defined to be the zero element of $\mathbf{R}[S]$). Then we may write any element of $\mathbf{R}[S]$ in the form $\sum_{p\in S} a_p e_p$. We can canonically associate any such expression with a unique *formal linear combination of elements of* S:

$$\sum_{p\in S} a_p\, e_p \longleftrightarrow \sum_{p\in S} a_p\, p. \tag{1.3}$$

It is conventional to write elements of the vector space $\mathbf{R}[S]$ in their associated "formal linear combination" notation, and to endow the set of formal linear combinations of elements of S with the vector-space structure induced by the 1-1 correspondence (1.3). It is easily seen that, in the "formal linear combination" notation for $\mathbf{R}[S]$, the vector-space operations are given by

$$\begin{aligned}\left(\sum_{p\in S} a_p\, e_p\right) + \left(\sum_{p\in S} b_p\, e_p\right) &= \sum_{p\in S} (a_p + b_p)\, e_p, \\ c\sum_{p\in S} a_p\, e_p &= \sum_{p\in S} (ca_p)\, e_p \quad \text{(for any } c\in\mathbf{R}).\end{aligned}$$

Since $1e_p = e_p$, it is also conventional to allow ourselves to write "$1p$" just as p for terms with coefficient 1 in a formal linear combination.

Remark 1.1 In the notation "$\{e_p\}_{p\in S}$," the set S is being used as an index-set for a natural basis $\mathcal{B}$ of $\mathbf{R}[S]$. This generalizes an example with which you are already familiar. If $S = \{1, 2, \ldots, n\}$, where $n \in \mathbf{N}$, then $\mathbf{R}[S]$ is canonically isomorphic to $\mathbf{R}^n$ (as is $\mathrm{Func}(S,\mathbf{R})$); an ordered n-tuple of real numbers is simply a presentation of a function $f : \{1, 2, \ldots, n\} \to \mathbf{R}$ as a tabulation of its values as a list $(f(1), f(2), \ldots, f(n))$. For this set S, the basis $\{e_p\}_{p\in S}$ is just the standard basis $\{e_j\}_{j=1}^n$ of $\mathbf{R}^n$.

[2] This is the same meaning of "free" as in "free module (over a ring)". A vector space is simply a module over a field.

For a general set S, formal-linear-combination notation simply takes advantage of the one-to-one correspondence $\mathcal{B} \leftrightarrow S$ (given by $e_p \leftrightarrow p$) to treat the set S notationally as *being* the basis $\mathcal{B}$, instead of just *indexing* $\mathcal{B}$. For the specific set $\{1, 2, \ldots, n\}$, we would never elect to use formal-linear-combination notation, however; the notation "$\sum_{j=1}^{n} a_j\, j$" (or "$\sum_{j\in\{1,\ldots n\}} a_j\, j$"), as opposed to would be too easily confused with the real number represented by the same notation, whereas the notation "$\sum_{j=1}^{n} a_j e_j$" does not have this problem. ▲

Remark 1.2 If the support of $f \in \mathbf{R}[S]$ is a given finite set $\{p_1, \ldots, p_N\}$, then instead of writing f using "Sigma notation", we may choose to write f in the form $a_1 e_{p_1} + \cdots + a_N e_{p_N}$. We allow ourselves to write the corresponding formal linear combination similarly, as $a_1 p_1 + \cdots + a_N p_N$. Of course, we also allow ourselves to write this as $\sum_{i=1}^{N} a_i p_i$. Similarly, if the support of f is a set $\{p_i\}$, with i running over a finite but unspecified set, we allow ourselves simply to use the notation $\sum_i a_i\, p_i$—with exceptions for some special a few sets S, as mentioned in Remark 1.1.

Remark 1.3 The expressions "$\sum_{p\in S} a_p\, p$" and "$a_1 p_1 + \cdots + a_N p_N$" are just *formal* linear combinations, since S here is just a *set*; in general there is no such thing as "2 times an element of S" or "the sum of two elements of S". It is important to keep in mind that even when S *does* have some algebraic structure (for example, if S itself is a vector space), *the space of formal linear combinations of elements of S does not change*. This space still means the vector space generated *freely* by the elements of S, and the expression "$\sum_{p\in S} a_p\, e_p$" is still just simplified notation for the finitely-supported function $p \mapsto a_p$ from S to $\mathbf{R}$.

Remark 1.4 The space of formal linear combinations also inherits a subtraction operation from $\mathbf{R}[S]$, which we still denote with a minus-sign. For example, if p, q are distinct elements of S, then then the formal linear combination $2p - 3q$ is just simplified notation for $2e_p - 3e_q = 2e_p + (-3)e_q$, the function $f : S \to \mathbf{R}$ satisfying $f(p) = 2, f(q) = -3$, and $f(x) = 0$ for other $x \in S$.

Remark 1.5 There is a natural injective map $\iota : S \to \mathbf{R}[S]$, namely $p \mapsto e_p$, which we sometimes refer to as the natural *inclusion* of S into $\mathbf{R}[S]$. Indeed, if we use formal-linear-combination notation, and allow ourselves to write "$1p$" simply as p then ι looks *exactly* like an inclusion map:

$$\begin{array}{rcl} \iota : S & \to & \mathbf{R}[S], \\ p & \mapsto & p. \end{array} \tag{1.4}$$

But take care not to be misled by the notational choices in (1.4): the p on the left of (1.4) is simply an element of the *set* S, which in general has no algebraic structure, whereas the p on the right of (1.4) is notation for a specific element of $\mathbf{R}[S]$ (namely e_p), a vector space no matter what the set S is!

Remark 1.6 Suppose that V is a nontrivial vector space and that $\mathcal{B}$ is a basis of V (see Section 6.1. Then every element of V can be written uniquely as $\sum_{v\in\mathcal{B}} a_v\, v$, where $a_v \in \mathbf{R}$ for all $v \in \mathcal{B}$, and $a_v = 0$ for all but finitely many $v \in \mathcal{B}$ (we again define the notation $\sum_{v\in\mathcal{B}} a_v\, v$ to mean the *finite* sum $\sum_{\{v\in\ \mathcal{B}\,|\,a_v\neq 0\}} a_v\, v$). Then the map

$$\begin{array}{rcl} V & \to & \mathbf{R}[\mathcal{B}], \\ \displaystyle\sum_{v\in\mathcal{B}} a_v\, v & \mapsto & \text{the } \textit{formal} \text{ linear combination } \displaystyle\sum_{v\in\mathcal{B}} a_v\, v, \end{array}$$

is an isomorphism. In other words, every nontrivial vector space V is isomorphic to the free vector space generated by any basis of V.

2 Tensor product of two vector spaces

Throughout this section, there is an implicit "Let V and W be vector spaces" whenever the symbols V and W occur in hypotheses, definitions, etc., and have not *explicitly* been assumed to be vector spaces.

2.1 Definitions

Some objects that will be central to our discussion are *bilinear* maps. These are a particular kind of map $V \times W \to Z$, where V, W, Z are vector spaces. For this discussion, it is *essential* that the reader put aside any habit of treating the notation "$V \times W$" as always meaning the vector space $V \oplus W$. When discussing bilinear maps, we definitely do *not* want to treat the set $V \times W$ as the vector space $V \oplus W$ (or as any other vector space).

A principle that should be clear by the end of this section is that *tensor product captures the essence of bilinearity.* This "essence" has nothing to do with finite-dimensionality, so we will *not* be assuming that our vector spaces are finite-dimensional.

Let V, W be vector spaces and let $\mathbf{R}[V \times W]$ be the free vector space generated by the *set* $V \times W$. We will shortly return to using the "formal linear combination" notation for elements of $\mathbf{R}[V\times W]$, but to avoid some potential confusion that Remark 1.3 already warned about) we will temporarily use our original function-notation for $\mathbf{R}[V \times W]$.

Define a set $\mathcal{R}(V,W) \subset \mathbf{R}[V \times W]$ by

$$\begin{array}{rcl} \mathcal{R}(V,W) & = & \{e_{(v_1+v_2,w)} - e_{(v_1,w)} - e_{(v_2,w)} \mid v_1, v_2 \in V; w \in W\} \\ & & \bigcup\{e_{(v,w_1+w_2)} - e_{(v,w_1)} - e_{(v,w_2)} \mid v \in V; w_1, w_2 \in W\} \\ & & \bigcup\{e_{(cv,w)} - ce_{(v,w)} \mid c \in \mathbf{R}; v \in V, w \in W\} \\ & & \bigcup\{e_{(v,cw)} - ce_{(v,w)} \mid c \in \mathbf{R}; v \in V, w \in W\}, \end{array}$$

and let $\mathcal{I}(V, W)$ denote the subpace of $\mathbf{R}[V \times W]$ generated by $\mathcal{R}(V, W)$. In "formal linear combination" notation,

$$
\begin{aligned}
\mathcal{R}(V, W) \quad = \quad & \{(v_1 + v_2, w) - (v_1, w) - (v_2, w) \mid v_1, v_2 \in V; w \in W\} & (2.5)\\
& \bigcup \{(v, w_1 + w_2) - (v, w_1) - (v, w_2) \mid v \in V; w_1, w_2 \in W\} & (2.6)\\
& \bigcup \{(cv, w) - c(v, w) \mid c \in \mathbf{R}; v \in V, w \in W\} & (2.7)\\
& \bigcup \{(v, cw) - c(v, w) \mid c \in \mathbf{R}; v \in V, w \in W\}, & (2.8)
\end{aligned}
$$

but in (2.5)–(2.8), it must be stressed that the minus-signs *do not* represent subtraction in the vector space $V \oplus W$, and that "$c(v, w)$" does *not* represent any operation in the space $V \oplus W$. The subtraction and scalar multiplication are operations on $\mathbf{R}[V \times W]$, not on $V \times W$. For example, the element of $\mathbf{R}[V \times W]$ denoted $c(v, w)$ (in formal-linear-combination notation) is a function $V \times W \to \mathbf{R}$ whose support consists of the single point (v, w), where the function takes the value 1. Unless $c = 1$, this function $c(v, w)$ is quite different from the element of $\mathbf{R}[V \times W]$ denoted (cv, cw), a function $V \times W \to \mathbf{R}$ whose support consists of the single point (cv, cw), where the function takes the value 1.

Definition 2.1 The vector space $V \otimes W$ (pronounced "V tensor W") is the quotient space $\mathbf{R}[V \times W]/\mathcal{I}(V, W)$.

The tensor-product symbol "$\otimes$" is also used on the level of *elements* $v \in V$ and $w \in W$:

Definition 2.2 For $v \in V, w \in W$, we define $v \otimes w \in V \otimes W$ by $v \otimes w = \pi(\iota(v, w))$, where $\iota : V \times W \to \mathbf{R}[V \times W]$ is the natural inclusion map (see Remark 1.5) and $\pi : \mathbf{R}[V \times W] \to \mathbf{R}[V \times W]/\mathcal{I}(V, W)$ is the quotient map. To avoid confusion in certain diagrams and expressions, we will sometimes write $\otimes_{\mathrm{op}}$ for the map $\pi \circ i : V \times W \to V \otimes W$ (so $v \otimes w = \otimes_{\mathrm{op}}(v, w)$ for $v \in V, w \in W$).

Thus $\pi(\sum_i c_i(v_i, w_i)) = \sum_i c_i(v_i \otimes w_i)$.

Proposition 2.3 *The following relations hold in $V \otimes W$:*

1. $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$ *for all* $v_1, v_2 \in V$ *and* $w \in W$.
2. $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$ *for all* $v \in V$ *and* $w_1, w_2 \in W$.
3. $(cv) \otimes w = c(v \otimes w) = v \otimes (cw)$ *for all* $v \in V, w \in W$ *and* $c \in \mathbf{R}$.

Proof: Exercise. ■

Corollary 2.4 *The map $\otimes_{\text{op}} : V \times W \to V \otimes W$ is bilinear, and its image spans $V \otimes W$.*

Proof: Exercise. ■

Remark 2.5 In general, the map $\otimes_{\text{op}} : V \times W \to V \otimes W$ is neither injective nor surjective, in general:

1. Two easy ways to see that injectivity fails are to observe that (i) for all nonzero $c \in \mathbf{R}$, we have $(cv) \otimes \frac{1}{c} w = v \otimes w$, and (ii) $0_V \otimes w = 0_{V \otimes W} = v \otimes 0_W$ for all $v \in V, w \in W$. Thus, injectivity fails unless $V = W = \{0\}$.

2. Surjectivity fails whenever $\dim(V)$ and $\dim(W)$ are at least 2. This follows quickly from a result proven later (Corollary 2.27) that shows that $V \otimes W$ has elements of rank 2. If $\otimes_{\text{op}}$ were surjective, then all elements of $V \otimes W$ would have rank 1 or 0.

Remark 2.6 The fact that the image of $\otimes_{\text{op}}$ spans $V \otimes W$ says precisely that every element of $V \otimes W$ can be written as a linear combination of elements of the form $v \otimes w$, i.e. as $\sum_i c_i(v_i \otimes w_i)$ for some finite lists of vectors $v_i \in V$, $w_i \in W$, and $c_i \in \mathbf{R}$. But given such lists, if define $v_i' = cv_i$, then $\sum_i c_i(v_i \otimes w_i) = \sum_i v_i' \otimes w_i$. Hence, more simply, we can write every element of $V \otimes W$ as a finite sum of the form $\sum_i v_i \otimes w_i$, i.e. a linear combination of elements of the form $v \otimes w$ but with all coefficients equal to 1.

For a given $T \in V \otimes W$, expressions of T in the form $\sum_i v_i \otimes w_i$ are *highly* non-unique, for several reasons. One reason is illustrated by the following example: suppose $v, v_1, v_2 \in V$ and $w, w_1, w_2 \in W$ are such that $v = v_1 + v_2$ and $w = w_1 + w_2$. Then

$$v \otimes w = v_1 \otimes w_1 + v_1 \otimes w_2 + v_2 \otimes w_1 + v_2 \otimes v_2 \ . \tag{2.9}$$

▲

The example (2.9) shows that we can always write any $T \in V \otimes W$ as an arbitrarily *long* sum of terms of the form $v \otimes w$. But a reasonable question, for a given T, is how *short* we can make such a sum.

Definition 2.7 Let $T \in V \otimes W$. We define the *rank* of T, as an element of $V \otimes W$ (there is an unrelated notion of *rank* that we will define in a later section), as follows:

- If $T = 0 = 0_{V \otimes W}$, the rank of T is $0 \in \mathbf{Z}$.

- If $T \neq 0$, the rank of T is the smallest number of summands in an expression of T as a finite sum of the form $\sum_i v_i \otimes w_i$. I.e. rank(T) is the smallest positive integer r such that
$$T = \sum_{i=1}^{r} v_i \otimes w_i \tag{2.10}$$
for some $v_1, \dots, v_r \in V$ and $w_1, \dots, w_r \in W$. We will refer to the right-hand side of (2.10) as a *minimal length expression* of a rank-r element of $V \otimes W$.

Exercise 2.8 Let $T \in V \otimes W$ be an element of rank $r > 0$, and let $\sum_{i=1}^{r} v_i \otimes w_i$ be a minimal-length expression of T. Show that each of the lists $(v_i)_{i=1}^{r}$ and $(w_i)_{i=1}^{r}$ is linearly independent.[4] (In particular, each of the sets $\{v_i : 1 \leq i \leq r\}$, $\{w_i : 1 \leq i \leq r\}$ is a linearly independent set with exactly r elements, so we can unambiguously use the notation $\{v_i\}_{i=1}^{r}$ and $\{w_i\}_{i=1}^{r}$ for these sets.) ▲

Even a minimal-length expression, as in (2.10), for a given element of $V \otimes W$ is nonunique. One issue is that for any nonzero $c \in \mathbf{R}$, we have $v \otimes w = (cv) \otimes \frac{1}{c} w$, so the element $v \otimes w \in V \otimes W$ does not even determine v and w uniquely. But, less trivially, suppose that $\{e_1, e_2\}$ and $\{f_1, f_2\}$ are linearly independent subsets of V and W, respectively. We will see later (Proposition 2.23) that $T := e_1 \otimes f_1 + e_2 \otimes f_2$ does not have rank 0 or 1, hence has rank 2. But having rank 2 does not determine e_1, e_2, f_1, f_2, even up to scalar multiples. For example, let $e_1' = e_1 + 3e_2$, $e_2' = 2e_1 + 4e_2$, $f_1' = -2f_1 + f_2$, and $f_2' = \frac{3}{2} f_1 - \frac{1}{2} f_2$. Then

$$\begin{aligned} e_1' \otimes f_1' + e_2' \otimes f_2' &= (e_1 + 3e_2) \otimes (-2f_1 + f_2) + (2e_1 + 4e_2) \otimes (\frac{3}{2} f_1 - \frac{1}{2} f_2) \\ &= e_1 \otimes f_1(-2+3) + e_1 \otimes f_2(1-1) + e_2 \otimes f_1(-6+6) \\ &\quad + e_2 \otimes f_2(3-2) \\ &= e_1 \otimes f_1 + e_2 \otimes f_2 \ . \end{aligned}$$

Exercise 2.9 Show that if either V or W is the trivial vector space $\{0\}$, then so is $V \otimes W$.

2.2 The universal property

Suppose now that V, W, Z are vector spaces and that $B : V \times W \to Z$ is a bilinear map. Can we define a *linear* map $L : V \otimes W \to Z$ by "setting $L(v \otimes w) = B(v, w)$ for all $(v, w) \in V \times W$ and extending linearly", analogously to the way we define linear maps by defining them on a basis and extending linearly? (We would wish, of course, for a *unique* such map L.)

[4] I used the word *lists* rather than *sets* or *ordered sets* initially since, by definition, the elements of a *set* are distinct, whereas a list can have "repeats".

It's not obvious that this procedure will work. Extending from a basis of a vector space relies on the fact that a basis is linearly independent. But the set $\{v \otimes w \mid v \in V, w \in W\}$—call this the set of *simple* elements of $V \otimes W$— is *highly* linearly dependent. If we can find a set of simple elements of $V \otimes W$ that is a basis of $V \otimes W$, we could define a linear map by defining it on this basis and extending linearly ... but there is never a *unique* such basis (the example leading to equation (2.9) illustrates part of the problem), and if $\dim(V)$ and $\dim(W)$ are at least 2, it is not obvious that the linear map we get is independent of the choice of simple basis.

If V and W are finite-dimensional, we can obtain a basis of simple elements, and show by "brute force" that the procedure above does determine a linear map $L : V \otimes W \to Z$ that satisfies $L(v \otimes w) = B(v, w)$ for all $(v, w) \in V \times W$ and is independent of the choice of basis. (You will do this for a related universal property later, in Exercise 4.20.) After doing things the hard way, we can better appreciate the value of the following proposition.

Proposition 2.10 (Universal Property of Tensor Products) *Let V, W be vector spaces.*

(a) The triple $(V \times W, V \otimes W, \otimes_{\mathrm{op}})$ has the following universal property: for any vector space Z, and any bilinear map $B : V \times W \to Z$, there exists a unique linear map $L_B : V \otimes W \to Z$ such that $L_B(v \otimes w) = B(v, w)$ for all $v \in V, w \in W$ (equivalently, such that $B = L_B \circ \otimes_{\mathrm{op}}$, as indicated by the commutative diagram in Figure 1).

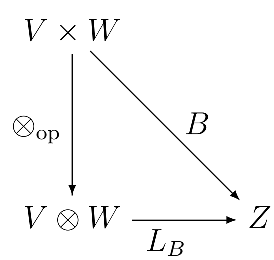

Figure 1: Diagram for Proposition 2.10(a)

(b) The pair $(\otimes_{\mathrm{op}}, V \otimes W)$ is "unique up to isomorphism", in the following sense: if X is a vector space and $T : V \times W \to X$ is another bilinear map such that $(V \times W, X, T)$ has the universal property described in part (a), then there is an isomorphism $L : V \otimes W \to X$ such that $T = L \circ \otimes_{\mathrm{op}}$.

Proof: (a) Let Z be a vector space and let $B : V \times W$ be a bilinear map.

Writing elements of $\mathbf{R}[V \times W]$ as formal linear combinations of elements of $V \times W$, the set $V \times W$ is a basis of $\mathbf{R}[V \times W]$. But *every* function from a basis of a vector space X to a vector space Y extends to a unique linear map from X to Y. In particular this is true of the function $B : V \times W \to Z$. Hence there is a unique linear map $\tilde{L} : \mathbf{R}[V \times W] \to Z$ such that $\tilde{L}(v, w) = B(v, w)$ for all $(v, w) \in V \times W$. Since $\tilde{L}$ is linear and B is bilinear, $\tilde{L}$ vanishes on every element of the set $\mathcal{I}(V, W)$ in (2.5)–(2.8), and hence vanishes on $\mathrm{span}(\mathcal{I}(V, W)) = \mathcal{I}(V, W) = \ker \pi$. Hence $\tilde{L}$ descends to a linear map $L : \mathbf{R}[V \times W]/\mathcal{I}(V, W) = V \otimes W \to Z$ satisfying $L(v \otimes w) = \tilde{L}(v, w) = B(v, w)$ for all $v \in V, w \in W$. Thus the diagram in Figure 1 commutes.

Suppose now that $L' : V \otimes W \to Z$ is another linear map for which this diagram commutes. Then $L' - L$ vanishes on the image of $\otimes_{\mathrm{op}}$. But this image spans $V \otimes W$. Since $L' - L$ is linear, it follows that $L' - L$ vanishes identically on $V \otimes W$. Hence $L' = L$, establishing the uniqueness asserted in the Proposition.

(b) Below, we refer to the universal property described in part (a) simply as "the universal property" for a given triple.

Let X be a vector space and let $T : V \times W \to X$ be a bilinear map such that $(V \times W, X, T)$ has the universal property. Then, since $\otimes_{\mathrm{op}}$ is bilinear, there exists a linear map $L_1 : X \to V \otimes W$ such that $\otimes_{\mathrm{op}} = L_1 \circ T$. Since T is bilinear, the universal property of $(V \times W, V \otimes W, \otimes_{\mathrm{op}})$ implies that there exists a linear map $L_2 : V \otimes W \to X$ such that $T = L_2 \circ \otimes_{\mathrm{op}}$. Hence $\otimes_{\mathrm{op}} = (L_1 \circ L_2) \circ \otimes_{\mathrm{op}}$. But, again applying the universal property of $(V \times W, V \otimes W, \otimes_{\mathrm{op}})$, this time with $Z = V \otimes W$ and $B = \otimes_{\mathrm{op}}$, there is a *unique* linear map $L : V \otimes W \to V \otimes W$ such that $\otimes_{\mathrm{op}} = L \circ \otimes_{\mathrm{op}}$. Since the identity map $\mathrm{id}_{V \otimes W}$ is a linear map $L : V \otimes W \to V \otimes W$ satisfying $\otimes_{\mathrm{op}} = L \circ \otimes_{\mathrm{op}}$, it follows that $L_1 \circ L_2 = I_{V \otimes W}$. Similarly, $L_2 \circ L_1 = \mathrm{id}_X$. Hence $L_2 : V \otimes W \to X$ is an isomorphism for which $T = L_2 \circ \otimes_{\mathrm{op}}$. ■

In view of Proposition 2.10, given a bilinear map $B : V \times W \to Z$, terminology such as "the linear map $L : V \times W \to Z$ induced by B" or "the linear map $L : V \otimes W \to Z$ determined by setting $L(v \otimes w) = B(v, w)$" is well-defined. Furthermore, if we wish to define a (particular) linear map $L : V \otimes W \to Z$, it suffices to define a bilinear map $B : V \times W \to Z$, and take L to be the linear map $V \otimes W \to Z$ induced by B.

Corollary 2.11 *Let V', W' be vector spaces and let $A : V \to V', B : W \to W'$ be linear maps. Then there is a unique linear map $L : V \otimes W \to V' \otimes W'$ satisfying $L(v \otimes w) = (Av) \otimes (Bw)$ for all $v \in V, w \in W$.*

Proof: Exercise. ■

Remark 2.12 The universal property in Proposition 2.10(a) is associated with the *triple* $(V \times W, V \otimes W, \otimes_{\mathrm{op}})$, not just the *pair* $(V \times W, V \otimes W)$, because this pair does noit uniquely determine a bilinear map $\otimes_{\mathrm{op}} : V \times W \to V \otimes W$ for which the stated universal property holds for $(V \times W, V \otimes W, \otimes'_{\mathrm{op}})$. For example, if we replace $\otimes_{\mathrm{op}}$ by $\otimes'_{\mathrm{op}} = 2\otimes_{\mathrm{op}}$, the triple $(V \times W, V \otimes W, \otimes_{\mathrm{op}})$ still has the universal property; each of the correspong linear maps $L'_B : V \times W \to Z$ is simply $\frac{1}{2}$ times the L_B in Figure1. More generally if $S : V \otimes W \to V \otimes V$ is any isomorphism, and we define $\otimes'_{\rm op} = S \circ \otimes_{\rm op}$, then $\otimes'_{\rm op} : V \times W \to V \otimes W$ is still bilinear, and the triple $(V \times W, V \otimes W, \otimes'_{\rm op})$ still has the universal property; each of the correspong linear maps $L'_B : V \times W \to Z$ is simply $L_B \circ S^{-1}$.

Exercise 2.13 [TO BE WRITTEN] What can't be changed in Prop 2.10:

The universal property in Proposition 2.10, rewritten in the equivalent way stated parenthetically, says:

> *[F]or any vector space Z, and any bilinear map $B : V \times W \to Z$, there exists a unique linear map $L_B : V \otimes W \to Z$ such that $B = L_B \circ \otimes_{\rm op}$.*

This is sometimes stated informally, as "every bilinear map $V \times W \to Z$ factors through a linear map $V \otimes W \to Z$," a *reminder* of much stronger statement made by the universal property. The informal statement is imprecise; all it says (literally) is that given any bilinear may $B : V \times W \to Z$, there exist (not necessarily unique) maps $f_B : V \times W \to V \otimes W$ (not necessarily bilinear) and a linear map $L_B : V \otimes W$, such that $B = g_B \circ f_B$. Even if strengthened to "every bilinear map $V \times W \to Z$ factors through a *unique* linear map $V \otimes W \to Z$," not strong enough.

With the informal statement interpreted literally, or with it replaced by several stronger statements that are still not as strong as

–linearity of L_B

–there isn't just *some* map $V \times W \to V \otimes W$ through which each B factors (with L_B linear); the map $\otimes_{\rm op}$ is *bilinear.*

–a bilinear map $B : W \to Z$ doesn't just factor through *some* bilinear map $V \times W \to V \otimes W$ and linear map $V \otimes W \to Z$; the bilinear map $\otimes_{\rm op}$ is the same for all B.

–If *uniqueness* of L_B deleted, then part (b) becomes false. (Related to: $\mathrm{im}(\otimes)_{\rm op}$ spans $V \otimes W$.)

Definition 2.14 (tensor product of two linear maps) Notation as in Corollary 2.11. The linear map L is called the *tensor product* of the linear maps A and B, and is denoted $A \otimes B$.

Thus, $A \otimes B : V \otimes W \to V' \otimes W'$ is the unique linear map satisfying

$$(A \otimes B)(v \otimes w) = Av \otimes Bw \tag{2.11}$$

for all $v \in V, w \in W$.

Observe that the map $\mathrm{id}_V \otimes \mathrm{id}_W : V \otimes W \to V \otimes W$ satisfies

$$(\mathrm{id}_V \otimes \mathrm{id}_W)(v \otimes w) = v \otimes w \quad \text{for all } v \in V,\ w \in W. \tag{2.12}$$

Since elements of the form $\{v \otimes w\}$ span $V \otimes W$, this implies that

$$\mathrm{id}_V \otimes \mathrm{id}_W = \mathrm{id}_{V \otimes W} \ . \tag{2.13}$$

Similarly, given vector spaces and maps $V \xrightarrow{A} V' \xrightarrow{C} V''$ and $W \xrightarrow{B} W' \xrightarrow{D} W''$, equation (3.22) yields

$$(C \otimes D)\big((A \otimes B)(v \otimes w)\big) = CAv \ \otimes \ DB\, w = (CA \otimes DB)(v \otimes w),$$

for all $v \in V,\ w \in W$, implying

$$(C \otimes D) \circ (A \otimes B) = (CA) \otimes (DB). \tag{2.14}$$

Exercise 2.15 Show that, for any vector space V, there are unique linear maps $V \otimes \mathbf{R} \to V$ and $\mathbf{R} \otimes V \to V$ satisfying $v \otimes 1 \mapsto v$ and $1 \otimes v \mapsto v$, respectively, and that these maps are isomorphisms.

Hence, $V \otimes \mathbf{R}$ and $\mathbf{R} \otimes V$ are canonically isomorphic to V. ▲

Remark 2.16 Let V, W be vector spaces and let $\tilde{\tau} : V \times W \to W \times V$ be defined by $\tilde{\tau}(v, w) = (w, v)$. Then $\tilde{\tau}$ is bilinear, and hence determines a linear "transpose" or "transposition" map $\tau : V \otimes W \to W \otimes V$ with the property that $\tau(v \otimes w) = w \otimes v$.

Exercise 2.17 Let V, W be vector spaces. Show that the transposition map $\tau : V \otimes W \to W \otimes V$ defined in Remark 2.16 is an isomorphism.

Remark 2.18 There are many isomorphisms that we consider to be *canonical*. The isomorphisms in Exercises 2.15 and 2.17 are among these. It is important to distinguish "canonically isomorphic" from "just-plain isomorphic", even though "canonically isomorphic" has no precise, universally applicable definition. For example, any two vector spaces V, W of the same finite dimension n are isomorphic, but, in general, are not *canonically* isomorphic (unless $n = 0$); there are inifinitely many isomorphisms from one space to the other, none singled out just by virtue of V and W being vector spaces of the same dimension.

In general, isomorphisms that we call canonical are induced by some defining structure of the spaces involved, and do not involve any *choices* not already made, e.g. choices of basis. Another rule of thumb is that whenever an isomorphism, or other map or structure, is called *canonical*, there are categories and functors lurking in the background.

Even among canonical isomorphisms, there are some that are "more canonical" than others, so much so that we treat the underlying spaces as beng *equal*, not just isomorphic. The isomorphisms in Exercise 2.15 are of this type; we allow ourselves to abuse notation and write simply "$V \otimes \mathbf{R} = V = \mathbf{R} \otimes V$," rather than something like "$V \otimes \mathbf{R} \underset{\text{canon.}}{\cong} V \underset{\text{canon.}}{\cong} \mathbf{R} \otimes V$." However, unless $W = V$, the transpose isomorphism in Exercise 2.17 is *not* of this type—we do not consider $V \otimes W$ and $W \otimes V$ to be the same space, for essentially the same reason that the *sets* $V \times W$ and $W \times V$ are not the same (even though there is a canonical bijection between them).

Remark 2.19 We have now used the symbol "$\otimes$", without any subscripts or superscripts, at three different "levels": at the level of vector spaces, where we have written $V \otimes W$; at the level of elements $v \in V, w \in W$, where we have written $v \otimes w$; and at the level of linear transformations $A : V \to V',\ B : W \to W'$, where we have written $A \otimes B$. The latter two uses are not quite on the same footing as the first use. If we let $\mathcal{V}$ denote the set of all vector spaces, then Definition 2.1 gives us a binary operation $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$,defined by $(V, W) \mapsto V \otimes W$. However, when we write "$v \otimes w$" (or even $\otimes_{\rm op}(v, w)$) as in Definition 2.2, the notation has no meaning *except* in the context of two given vector spaces V and W, of which v and w (respectively) are elements. We have *not* defined a binary operation on $\bigcup\{V : V \text{ is a vector space}\}$; in the expression $v \otimes w$, "$\otimes$" is not a binary operation on one universal set. Notation such as "$v \otimes^{(V,W)} w$" and "$\otimes_{\rm op}^{(V,W)}(v, w)$" would be more precise than $v \otimes w$ or $\otimes_{\rm op}(V, W)$. Similarly, when we write "$A \otimes B$" for linear transformations A, B, the notation makes sense only after specifying the domains and codomains of A and B; we have not introduced a binary operation on the set of *all* linear transformations.

The cleanest way to assemble these different uses of "$\otimes$" into a coherent whole is via appropriate categories and functors; see Remark 2.20 and Exercise 2.21 below.

Remark 2.20 Let $\mathcal{C}$ be the category whose objects are vector spaces and whose morphisms linear transformations. Then the tensor-product operations we have defined for vector spaces and linear transformations can be encoded as a (covariant) functor "$\widehat{\bigotimes}$" from the product category $\mathcal{C} \times \mathcal{C}$ to the category $\mathcal{C}$. In $\mathcal{C} \times \mathcal{C}$, the objects are ordered pairs (V, W), where V and W are vectors spaces, and a morphism from (V, W) to (V', W') is a pair of linear maps $(A, B) \in \text{Hom}(V, V') \times \text{Hom}(W, W')$. For this functor, the map from objects to objects is defined by setting

$$\widehat{\bigotimes}\big(\,(V, W)\,\big) = V \otimes W.$$

For objects $(V, W), (V', W')$ of $\mathcal{C} \times \mathcal{C}$, the map

$$\underbrace{\text{Mor}\big(\,(V, W),\ (V', W')\,\big)}_{\text{morphisms in } \mathcal{C}\times\mathcal{C}} \to \underbrace{\text{Mor}(V \otimes W,\ V' \otimes W')}_{\text{morphisms in } \mathcal{C}} = \text{Mor}(\widehat{\bigotimes}\big(\,(V, W)\,\big), \widehat{\bigotimes}\big(\,(V', W')\,\big)$$

is defined by

$$\widehat{\bigotimes}\big(\,(A, B)\,\big) = A \otimes B$$

(where the RHS is defined as in Definition 2.14). Noting that the identity morphism of an object (V, W) of $\mathcal{C} \times \mathcal{C}$ is $(\text{id}_V, \text{id}_W)$ (by definition of "product category"), equations (2.13) and (2.14), taken together, are precisely the statement that $\widehat{\bigotimes}$ is a covariant functor.

Note that the statement "$\widehat{\bigotimes}$ is a functor" does not tell us how to *define* the tensor product of vector spaces, linear transformations, or elements of vector spaces; it simply encodes certain relationships. Having *previously defined* the element-level maps "$\otimes_{\text{op}}^{(V,W)}$" : $V \times W \to V \otimes W$, we were able to use them Definition 2.14 to define the tensor product of linear transformations, from which the morphism-requirement equations (2.12) and then (2.13) followed. But if we *start* with the morphism-requirement (2.13), we cannot deduce what the element "$v \otimes w$" of $V \otimes W$ is from equation (2.12). The functor $\widehat{\bigotimes}$, as defined above, encodes tensor-product relations for *vector spaces* (as objects, not as sets) and *linear transformations*, but not for *elements* of vector spaces.

To obtain a functor $\bigotimes$ that encodes, additionally, the relation between $\widehat{\bigotimes}$ and the *element-level* maps "$\otimes_{\text{op}}^{(V,W)}$" : $V \times W \to V \otimes W$", we need to start with a category more refined than $\mathcal{C}$. See Exercise 2.21 below. ▲

Exercise 2.21 Let $\mathcal{C}'$ be the category whose objects are pairs (V, v), where V is a vector space and $v \in V$, and whose morphisms from (V, v) to (W, w) are linear transformations carrying v to w:

$$\text{Mor}\big((V, v),\ (W, w) \big) = \{A \in \text{Hom}(V, W) : Lv = w\}.$$

(a) Show that $\mathcal{C}'$ is, indeed, a category, where we define the composition of morphisms to be the composition of the corresponding linear maps.

(b) Given an object $\big((V, v),\ (W, w) \big)$ of $\mathcal{C}' \times \mathcal{C}'$, define an object $\bigotimes \big((V, v),\ (W, w) \big)$ of $\mathcal{C}'$ by

$$\bigotimes \Big(\big((V, v),\ (W, w) \big) \Big) = (V \otimes W,\ v \otimes w)$$

Given objects $\big((V, v),\ (W, w)\big)$ and $\big((V', v'),\ (W', w')\big)$ of $\mathcal{C}' \times \mathcal{C}'$, and a $(\mathcal{C}' \times \mathcal{C}')$-morphism $(A, B) \in \text{Mor}\Big(\big((V, v),\ (W, w)\big),\ \big((V', v'),\ (W', w')\big) \Big)$, define a $\mathcal{C}'$ morphism

$$\bigotimes \big((A, B) \big) \quad \in \quad \text{Mor}\big((V \otimes W, v \otimes w),\ (V' \otimes W', v' \otimes w') \big)$$

by

$$\bigotimes \big((A, B) \big) = A \otimes B.$$

Show that $\bigotimes$ is a covariant functor from $\mathcal{C}' \times \mathcal{C}'$ to $\mathcal{C}'$. ▲

2.3 Bases of $V \otimes W$

We start with an exercise that will be used in the proof of a subsequent proposition:

Exercise 2.22 Let $(v, w) \in V \times W$. Show that if $v \otimes w = 0$ then $v = 0$ or $w = 0$.

Proposition 2.23 *Let $\mathcal{A}_V$ and $\mathcal{A}_W$ be linearly independent subsets of V and W respectively. Then the map $\mathcal{A}_V \times \mathcal{A}_W \to V \otimes W$, $(v, w) \mapsto v \otimes w$, is one-to-one, and the set $\{v \otimes w \mid v \in \mathcal{A}_V, w \in \mathcal{A}_W\}$ is linearly independent.*

Proof: It suffices to show that given any distinct $e_1, \ldots, e_n \in \mathcal{A}_V$ and distinct $f_1, \ldots, f_m \in \mathcal{A}_W$, all the elements $e_i \otimes f_j$ are distinct, and the set

$$\{e_i \otimes f_j : i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}\} \tag{2.15}$$

is linearly independent.

First, suppose there are distinct index-pairs $(i_1, j_1), (i_2, j_2)$ such that

$$e_{i_1} \otimes f_{j_1} = e_{i_2} \otimes f_{j_2} \tag{2.16}$$

Without loss of generality we may assume that $(i_1, j_1) = (1, 1)$ and that (i_2, j_2) is one of the pairs $(2, 1), (1, 2)$, or $(2, 2)$.

If $(i_2, j_2) = (2, 1)$, then (2.16) becomes $e_1 \otimes f_1 = e_2 \otimes f_1$, implying $(e_1 - e_2) \otimes f_1 = 0$. Hence, by Exercise 2.22, at least one of the vectors $e_1 - e_2$ and f_1 must be zero. Since the e_i are assumed distinct, $e_1 - e_2 \neq 0$, so $f_1 = 0$, contradicting the assumed linear independence of $\{f_i\}_{i=1}^m$. Hence $(i_2, j_2) \neq (2, 1)$.

Similarly $(i_2, j_2) \neq (1, 2)$. Hence $(i_2, j_2) = (2, 2)$, and (2.16) becomes $e_1 \otimes f_1 = e_2 \otimes f_2$.

Extend the linearly independent sets $\{e_1, e_2\}$ and $\{f_1, f_2\}$ to bases $\mathcal{B}_V, \mathcal{B}_W$ of V, W respectively (see Remark 6.1). Let $\theta : V \to \mathbf{R}$ and $\varphi : W \to \mathbf{R}$ be the unique linear maps (hence elements of V^*, W^* respectively) satisfying

$$\begin{aligned} \theta(e_1) = 1 &= \varphi(f_1), \\ \theta(v) = 0 &\text{ if } v \in \mathcal{B}_V \setminus \{e_1\}, \\ \varphi(w) = 0 &\text{ if } w \in \mathcal{B}_W \setminus \{f_1\}. \end{aligned}$$

Define $B : V \times W \to \mathbf{R}$ by $B(v', w') = \langle \theta, v' \rangle \langle \varphi, w' \rangle$. (Here and below, the notation $\langle \cdot, \cdot \rangle$ is being used for both the dual pairings $V^* \times V \to \mathbf{R}$ and $W^* \times W \to \mathbf{R}$; see Section 6.2.) Then B is bilinear, hence determines a linear map $L : V \otimes W \to \mathbf{R}$ satisfying $L(v' \otimes w') = \langle \theta, v' \rangle \langle \varphi, \omega' \rangle$ for all $(v', w') \in V \times W'$. But then $L(e_1 \otimes f_1) = B(e_1, f_1) = 1$ while $L(e_2 \otimes f_2) = B(e_2, f_2) = 0$, contradicting $e_1 \otimes f_1 = e_2 \otimes f_2$.

Hence $e_{i_1} \otimes f_{j_1} \neq e_{i_2} \otimes f_{j_2}$ whenever $(i_1, j_1) \neq (i_2, j_2)$.

Suppose now that $\{c^{ij}\}_{i \in \{1, \ldots, n\}, j \in \{1, \ldots m\}}$ is a collection of real numbers for which $\sum_{i,j} c^{ij} \, e_i \otimes f_j = 0$. Extend $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ to bases $\mathcal{B}_V, \mathcal{B}_W$ of V, W respectively. For $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$ define $\theta^i \in V^*$, $\varphi^j \in W^*$ to be the unique linear maps satisfying

$$\begin{aligned} \theta^i(v) &= \begin{cases} 1 & \text{if } v = e_i, \\ 0 & \text{if } v \in \mathcal{B}_V \setminus \{e_i\}, \end{cases} \\ \varphi^j(w) &= \begin{cases} 1 & \text{if } w = f_j, \\ 0 & \text{if } w \in \mathcal{B}_W \setminus \{f_j\}, \end{cases} \end{aligned}$$

For $1 \le i \le n$ and $1 \le j \le m$ let $B^{(i,j)} : V \times W \to \mathbf{R}$ be the blinear map defined by $B^{(i,j)}(v,w) = \langle \theta^i, v\rangle \langle \varphi^j, w\rangle$, and let $L^{(i,j)} : V \otimes W \to \mathbf{R}$ be the induced linear map. Then for all (i,j) we have

$$0 = L^{(i,j)}(0) = L^{(i,j)}\left(\sum_{k,l} c^{kl}\, e_k \otimes f_l\right) = \sum_{k,l} c^{kl} B^{(i,j)}(e_k, f_l) = \sum_{k,l} c^{kl}\, \delta^i_k\, \delta^j_l = c^{ij}.$$

Hence the set (2.15) is linearly independent. ■

Corollary 2.24 *Let $\mathcal{B}_V, \mathcal{B}_W$ be bases of the vector spaces V, W respectively. Then the set*

$$\{v \otimes w \mid v \in \mathcal{B}_V,\ w \in \mathcal{B}_W\} \tag{2.17}$$

is a basis of $V \otimes W$.

Proof: Suppose $T \in V \otimes W$. Then $T = \sum_{i=1}^r v_i \otimes w_i$ for some $v_1, \dots, v_r \in V$ and $w_1, \dots w_r \in W$. Expanding each v_i (respectively w_i) in terms of the basis $\mathcal{B}_V$ (resp. $\mathcal{B}_W$), and using bilinearity of $\otimes_{\text{op}}$, we obtain a re-expression of T as sum of the form $\sum_{v \in \mathcal{B}_V, w \in \mathcal{B}_W} c_{(v,w)} v \otimes w$, where all but finitely many of the coefficients $c_{(v,w)}$ are zero. Hence the set (3.23) spans $V \otimes W$. By Proposition 2.23, this set is also linearly independent, hence is a basis of $V \otimes W$. ■

2.4 Several canonical maps

For any vector spaces V and W, the set of bilinear maps $V \times W \to \mathbf{R}$ is easily seen to be a subspace of $\mathbf{R}^{V \times W}$. We denote this subspace $\text{Bihom}(V \times W, \mathbf{R})$ and call it the *space* of bilinear maps $V \times W \to \mathbf{R}$.

As we will see in Section 2.5, *when V and W are finite-dimensional*, several spaces built from V and W are canonically isomorphic. For example, $V^* \otimes W^* \underset{\text{canon.}}{\cong} \text{Bihom}(V \times W, \mathbf{R})$ (which is often taken as a "quick and dirty" definition of $V^* \otimes W^*$ in the finite-dimensional case); $V^* \otimes W \underset{\text{canon.}}{\cong} \text{Hom}(V, W)$; and $V^* \otimes W^* \underset{\text{canon.}}{\cong} (V \otimes W)^*$. These finite-dimensional isomorphisms stem from canonical *linear maps* from the space on the left of the "$\underset{\text{canon.}}{\cong}$" symbol to the space on the right, maps whose definitions do not require finite-dimensionality. In general, without the finite-dimensionality assumption, none of these maps is an isomorphism, but they do turn out to be canonical *injections* (which can be thought of as inclusion maps). There is also a canonical isomorphism not listed above: $(V \otimes W)^* \underset{\text{canon.}}{\cong} \text{Bihom}(V \times W, \mathbf{R})$.

In the proposition below, the canonical linear maps mentioned above are defined (for arbitrary vector spaces), and their injectivity (and, in one case, surjectivity) is established. We will make frequent use of the universal property of tensor products (Proposition 2.10) without explicit reference to the term "universal property".

Proposition 2.25 *Let V, W be vector spaces.*

(a) *For each $B \in \mathrm{Bihom}(V \times W, \mathbf{R})$, let $L_B \in (V \otimes W)^*$ be the linear map $V \otimes W \to \mathbf{R}$ induced by B. Then the map*

$$\begin{aligned} \iota : \mathrm{Bihom}(V \times W, \mathbf{R}) &\to (V \otimes W)^*, \\ B &\mapsto L_B, \end{aligned}$$

is an isomorphism. Hence $(V \otimes W)^*$ is canonically isomorphic to $\mathrm{Bihom}(V \times W, \mathbf{R})$.

(b) *For each $(\theta, \varphi) \in V^* \times W^*$ define a map $B_{(\theta,\varphi)} : V \times W \to \mathbf{R}$ by*

$$B_{(\theta,\varphi)}(v, w) = \langle \theta, v \rangle \langle \varphi, w \rangle. \tag{2.18}$$

(Here the notation $\langle \cdot, \cdot \rangle$ is being used for both the dual pairings $V^* \times V \to \mathbf{R}$ and $W^* \times W \to \mathbf{R}$; see Section 6.2.) *Then:*

(i) *For each $(\theta, \varphi) \in V^* \times W^*$, the map $B_{(\theta,\varphi)} : V \times W \to \mathbf{R}$ is bilinear, hence an element of $\mathrm{Bihom}(V \times W, \mathbf{R})$.*

(ii) *The map $V^* \times W^* \to \mathrm{Bihom}(V \times W, \mathbf{R})$ given by $(\theta, \varphi) \mapsto B_{(\theta,\varphi)}$ is bilinear, and hence induces a linear map*

$$\begin{aligned} j : V^* \otimes W^* &\to \mathrm{Bihom}(V \times W, \mathbf{R}) \\ \text{satisfying} \quad j(\theta \otimes \varphi) &= B_{(\theta,\varphi)} \quad \text{for all } (\theta, \varphi) \in V^* \otimes W^*. \end{aligned}$$

(iii) *The linear map j is injective.*

Hence there is a canonical injection

$$j : V^* \otimes W^* \hookrightarrow \mathrm{Bihom}(V \times W, \mathbf{R}).$$

(c) *There is a canonical injection*

$$V^* \otimes W^* \hookrightarrow (V \otimes W)^*,$$

namely the map $\iota \circ j$.

(d) *For each $(w, \theta) \in W \times V^*$, define $T_{(w,\theta)} : V \to W$ by*

$$T_{(w,\theta)}(v) = \langle \theta, v \rangle w. \tag{2.19}$$

Then:

(i) *For each $(w, \theta) \in W \times V^*$, the map $T_{(w,\theta)} : W \times V^* \to \mathbf{R}$ is linear, hence an element of $\mathrm{Hom}(V, W)$.*

(ii) *The map* $W \times V^* \to \mathrm{Hom}(V, W)$ *given by* $(w, \theta) \mapsto T_{(w,\theta)}$ *is bilinear, and hence induces a linear map*

$$\begin{array}{rrcl} & \hat{\jmath} : W \otimes V^* & \to & \mathrm{Hom}(V, W) \\ \text{satisfying} & \hat{\jmath}(w \otimes \theta) & = & T_{(w,\theta)} \quad \text{for all } (w, \theta) \in W \otimes V^*. \end{array}$$

(iii) *The linear map* $\hat{\jmath}$ *is injective.*

Hence, letting $\tau : V^* \otimes W \to W \otimes V^*$ *denote the transposition isomorphism* (see Remark 2.16 and Exercise 2.17), *the maps*

$$\begin{array}{rrcl} & \hat{\jmath} : W \otimes V^* & \hookrightarrow & \mathrm{Hom}(V, W) \\ \text{and} & \hat{\jmath} \circ \tau : V^* \otimes W & \hookrightarrow & \mathrm{Hom}(V, W) \end{array}$$

are canonical injections.

(iv) *Let* $T \in W \otimes V^*$, *let* $r = \mathrm{rank}(T)$ *be the rank of* T *as an element of* $W \otimes V^*$ (see Definition 2.7), *and, if* $r > 0$, *let* $\sum_{i=1}^r w_i \otimes \xi^i$ *be a minimal-length expression of* T. *Then if* $r > 0$, $\mathrm{im}(\hat{\jmath}(T)) = \mathrm{span}\{w_i : 1 \le i \le r\}$, *while if* $r = 0$ *then* $\mathrm{im}(\hat{\jmath}(T)) = \{0\}$. *It follows that the rank of* T *as an element of* $W \otimes V^*$ *coincides with the rank of the linear transformation* $\hat{\jmath} : V \to W$ *(i.e. the dimension of* $\mathrm{im}(\hat{\jmath}(T))$*).*

(e) *For each* $(v, w) \in V \times W$, *define maps* $T'_{(v,w)} : V^* \to W$ *and* $T''_{(v,w)} : W^* \to V$ *by* $T'_{(v,w)}(\theta) = \langle \theta, v \rangle w$ (i.e. $T'_{(w,v)}(\theta) = T_{(w,\theta)}(v)$, where $T_{(w,\theta)}$ is as in (2.19)) *and, analogously,* $T''_{(v,w)}(\varphi) = \langle \varphi, w \rangle v$. *Then:*

(i) *For each* $(v, w) \in V \times W$, *the maps* $T'_{(v,w)} : V^* \to W$ *and* $T''_{(v,w)} : W^* \to V$ *are linear, hence are elements of* $\mathrm{Hom}(V^*, W)$ *and* $\mathrm{Hom}(W^*, V)$ *respectively.*

(ii) *The maps* $T' : V \times W \to \mathrm{Hom}(V^*, W)$ *and* $T'' : V \times W \to \mathrm{Hom}(W^*, V)$ *defined by* $T'(v, w) = T'_{(v,w)}$ *and* $T''(v, w) = T''_{(v,w)}$ *are bilinear, and hence induce linear maps*

$$\begin{array}{rcl} j' : V \otimes W & \to & \mathrm{Hom}(V^*, W), \\ j'' : V \otimes W & \to & \mathrm{Hom}(W^*, V) \end{array}$$

satisfying $j'(v \otimes w) = T'_{(v,w)}$ *and* $j''(v \otimes w) = T''_{(v,w)}$ *for all* $(v, w) \in V \otimes W$.

(iii) *The linear maps* j' *and* j'' *are injective.*

(iv) *Let* $T \in V \otimes W$, *let* $r = \mathrm{rank}(T)$ *be the rank of* T *as an element of* $V \otimes W$ (see Definition 2.7), *and, if* $r > 0$, *let* $\sum_{i=1}^r v_i \otimes w_i$ *be a minimal-length expression of* T. *Then if* $r > 0$, $\mathrm{im}(j'(T)) = \mathrm{span}\{w_i : 1 \le i \le r\})$ *and* $\mathrm{im}(j''(T)) = \mathrm{span}\{v_i : 1 \le i \le r\}$, *while if* $r = 0$ *then* $\mathrm{im}(j'(T)) = \{0_W\}$ *and* $\mathrm{im}(j''(T)) = \{0_V\}$. *It follows that the rank of* T *as an element of* $V \otimes W$, *and the ranks of the linear transformations* $j'(T) : V^* \to W$ *and* $j''(T) : W^* \to V$, *all coincide.*

(v) *Let $r > 0$ and suppose that $\{v_1, \ldots, v_r\}$ and $\{w_1, \ldots w_r\}$ are linearly independent subsets of V and W respectively. Let $T = \sum_{i=1}^r v_i \otimes w_i$. Then* $\mathrm{rank}(T) = r$.

Remark 2.26 In the finite-dimensional case, $V^{**} = V$, so part (e) would essentially be redundant with part (d). But since we are not assuming finite-dimensionality of V or W in Proposition 2.25, part (e) needs its own statement and proof.

Proof of Proposition 2.25: (a) Observe that if $B_1, B_2 \in \mathrm{Bihom}(V \times W, \mathbf{R})$ and $c_1, c_2 \in \mathbf{R}$, then for all $(v, w) \in V \times W$ we have

$$\begin{aligned} L_{c_1B_1+c_2B_2}(v \otimes w) = (c_1B_1 + c_2B_2)(v, w) &= c_1B_1(v, w) + c_2B_2(v, w) \\ &= c_1L_{B_1}(v \otimes w) + c_2L_{B_2}(v \otimes w). \end{aligned}$$

It follows that $L_{c_1B_1+c_2B_2} = c_1L_{B_1} + c_2L_{B_2}$. It follows that the map $\iota : \mathrm{Bihom}(V \times W, \mathbf{R}) \to (V \otimes W)^*$ defined by $\iota(B) = L_B$ is linear. We claim that ι is an isomorphism.

First, suppose $B \in \ker(\iota)$. Then $L_B = 0$, so $B(v, w) = L_B(v \otimes w) = 0$ for all $(v, w) \in V \times W$; i.e. $B = 0$. Hence ι is injective.

To establish surjectivity, let $T \in (V \otimes W)^*$. Define $B : V \times W \to \mathbf{R}$ by $B(v, w) = T(v \otimes w)$. Then $T = L_B = \iota(B)$. Hence ι is surjective, and is therefore an isomorphism.

(b) The bilinearity asserted in statements (i) and (ii) are easily seen from (2.18). The remainder of (ii) follows from Proposition 2.10. For (iii), suppose $\ker L \neq 0$ and let T be a nonzero element of $\ker(L)$. Express T as a minimal-length sum of simple elements of $V^* \otimes W^*$:

$$T = \sum_{i=1}^r \xi^i \otimes \eta^i ,$$

where $r = \mathrm{rank}(T) \geq 1$. By Exercise 2.8, $\{\xi^i\}_{i=1}^r$ and $\{\eta^i\}_{i=1}^r$ are linearly independent sets in V^*, W^* respectively. Hence there exist sets $\{v_i\}_{i=1}^r$, $\{w_i\}_{i=1}^r$ in V, W respectively, such that $\langle \xi^i, v_j \rangle = \delta^i_j = \langle \varphi^i, w_j \rangle$ for all $i, j \in \{1, \ldots, r\}$ (see Exercise 6.17). Since $L(T) = 0$, we have

$$0 = L(T)(v_1, w_1) = \left(\sum_{i=1}^r L(\xi^i \otimes \eta^i) \right) (v_1, w_1) = \sum_{i=1}^r \langle \xi^i, v_1 \rangle \langle \eta^i, w_1 \rangle = \sum_{i=1}^r \delta_{i1}\delta_{i1} = 1,$$

a contradiction.

Hence $\ker(T) = \{0\}$, and L is injective.

(c) Immediate from (a) and (b).

(d) Items (i)–(iii): exercise.

(iv) First suppose $r = 0$. Then $T = 0$, so $\hat{j}(T) = 0 \in \mathrm{Hom}(V, W)$, implying $\mathrm{im}(\hat{j}(T)) = \{0\}$ and $\mathrm{rank}(\hat{j}(T)) = 0 = r$.

Now suppose $r > 0$, and let $\sum_{i=1}^r w_i \otimes \xi^i$ be a minimal-length expression of T. From the definition of $\hat{j}(T)$, it is clear that $\mathrm{im}(\hat{j}(T)) \subset \mathrm{span}\{w_1, \ldots, w_r\}$ (this would be true even without the "minimal length" hypothesis). As in the proof of part (b), because "minimal length" ensures that the set $\{\xi^i\}_{i=1}^r$ is linearly independent (Exercise 2.8), we may choose $v_1, \ldots, v_r \in V$ such that $\langle \xi^i, v_j \rangle = \delta^i_j = \langle \varphi^i, w_j \rangle$ for all $i, j \in \{1, \ldots, r\}$. But then for each $j \in \{1, \ldots, r\}$,

$$\hat{j}(T)(v_j) \;=\; \sum_{i=1}^r T_{(w_i, \xi^i)}(v_j) \;=\; \sum_{i=1}^r \langle \xi^i, v_j \rangle w_i \;=\; w_j \;.$$

Hence $\mathrm{span}\{w_1, \ldots, w_r\} \subset \mathrm{im}(\hat{j}(T))$. Thus $\mathrm{im}(\hat{j}(T)) = \mathrm{span}\{w_i : 1 \leq i \leq r\}$. But, using Exercise 2.8 again, the set $\{w_i\}_{i=1}^r$ is linearly independent, so $\mathrm{rank}(\hat{j}(T)) = \dim(\mathrm{span}\{w_i\}_{i=1}^r) = r = \mathrm{rank}(T)$.

(e) Exercise. ■

Corollary 2.27 *Let $r > 0$ and suppose that* $\dim(V)$ *and* $\dim(W)$ *are both at least r. Then $V \otimes W$ contains elements of rank r.*

Proof: This is immediate from Proposition 2.25(e)(v). ■

Corollary 2.28 *Assume that at least one of the spaces V, W is finite-dimensional. Then the* maximal rank *of elements of $V \otimes W$ is* $\min\{\dim(V), \dim(W)\}$ *. I.e. there exists $T \in V \otimes W$ for which* $\mathrm{rank}(T) = \min\{\dim(V), \dim(W)\}$*, and no element of $V \otimes W$ has larger rank.*

Proof: Let $r = \min\{\dim(V), \dim(W)\}$. If $r = 0$ then at least one of the vector spaces V, W is $\{0\}$, so $V \otimes W = \{0\}$ and there is nothing to prove.

Thus, assume $r > 0$. Corollary 2.27 shows that $V \otimes W$ has an element of rank r. Exercise 2.8 shows that no element of $V \otimes W$ can have rank greater than r. ■

2.5 The finite-dimensional case

In this subsection, we assume that both V and W are finite-dimensional, and let $n = \dim(V)$, $m = \dim(W)$. We will assume $n, m \geq 1$, since the cases $n = 0$ and $m = 0$ are uninteresting (everything that needs to be said about these cases follows immediately from Exercise 2.9).

Proposition 2.29 *Let* $\{e_i\}_{i=1}^n$ *and* $\{f_i\}_{i=1}^m$ *be bases of* V *and* W *respectively. Then*

$$\{e_i \otimes f_j\}_{1\le i\le n,\ 1\le j\le m}$$

is basis of $V \otimes W$, *and hence*

$$\dim(V \otimes W) = \dim(V)\,\dim(W).$$

Proof: Immediate from Corollary 2.24. ■

Proposition 2.30 *Let* $\iota : \mathrm{Bihom}(V \times W, \mathbf{R}) \to (V \otimes W)^*$ *be the canonical isomorphism defined in Proposition 2.25(a), let* $j : V^* \otimes W^* \to \mathrm{Bihom}(V \times W, \mathbf{R})$ *and* $\hat{j} : W \otimes V^* \to \mathrm{Hom}(V, W)$ *be the canonical injections defined in Proposition 2.25 parts (b) and (c), and let* $\tau : V^* \otimes W \to W \otimes V^*$ *be the transposition isomorphism* (see Remark 2.16 and Exercise 2.17)*. Then all of the following maps are isomorphisms:*

$$\begin{aligned} j : V^* \otimes W^* &\to \mathrm{Bihom}(V \times W, \mathbf{R}), \\ \iota \circ j : V^* \otimes W^* &\to (V \otimes W)^*, \\ \hat{j} : W \otimes V^* &\to \mathrm{Hom}(V, W), \\ \text{and} \qquad \hat{j} \circ \tau : V^* \otimes W &\to \mathrm{Hom}(V, W). \end{aligned}$$

Hence, in each of the four lines above, the domain and codomain of the indicated map are canonically isomorphic.

Proof: We have already seen that all four of the linear maps $j,\ \iota \circ j,\ \hat{j}$, and $\hat{j} \circ \tau$ are injections. By finite-dimensionality, $\dim(V^*) = \dim(V) = n$ and $\dim(W^*) = \dim(W) = m$, so (by Proposition 2.29) the spaces $V \otimes W$, $(V \otimes W)^*$, $V^* \otimes W^*$, $W \otimes V^*$, and $V^* \otimes W$ all have dimension nm. Since ι is an isomorphism, $\mathrm{Bihom}(V \times W, \mathbf{R})$ also has dimension nm. From elementary linear algebra, $\mathrm{Hom}(V, W)$ is (non-canonically) isomorphic to the space of real $m \times n$ matrices, hence has dimension nm as well.

Hence, by equidimensionality of domain and codomain, the canonical linear injections $j,\ i \circ j,\ \hat{j}$, and $\hat{j} \circ \tau$, are all isomorphisms. ■

Remark 2.31 ("tensor product" of two matrices) Let V', W' be vector spaces of dimension $n', m' \ge 1$, and let $A \in \mathrm{Hom}(V, V'), B \in \mathrm{Hom}(W, W')$. Keeping in mind that matrix of a linear map $V \otimes W \to V' \times W'$, with respect to any given bases of $V \otimes W$ and $V' \otimes W'$, will be of size $(n'm') \times (nm)$, what is a good way to express the matrix of $A \otimes B$ with respect to bases of the form in Proposition 2.29?

Let $\mathbf{e} := \{e_i\}_{i=1}^n$, $\mathbf{f} := \{f_i\}_{i=1}^m$, $\mathbf{e}' := \{e_i'\}_{i=1}^{n'}$, $\mathbf{f}' := \{f_i'\}_{i=1}^{m'}$ be bases of V, W, V', W' respectively, and let $\tilde{A}$ be the matrix of A with respect to the bases $\mathbf{e}$ and $\mathbf{e}'$, and let $\tilde{B}$ be the matrix of B with respect to the bases $\mathbf{f}$ and $\mathbf{f}'$. Then, using Einstein summation convention (i.e., summing over any repeated index that occurs once upstairs and once downstairs), $Ae_j = e_i \tilde{A}^i{}_j$ and $Bf_l = f_k \tilde{B}^k{}_l$, so

$$(A \otimes B)(e_j \otimes f_l) = e'_i \otimes f'_k \ \tilde{A}^i{}_j \tilde{B}^k{}_l \tag{2.20}$$

(note that we are summing over both i and k on the RHS of (2.20)). To express this in terms of matrix of the appropriate size, we need to *order* the bases $\{e_j \otimes f_l\}$ and $\{e'_i \otimes f'_k\}$. For each of these, we choose the lexicographic ordering:

$$e_1 \otimes f_1,\ e_1 \otimes f_2, \dots, e_1 \otimes f_m,\ e_2 \otimes f_1,\ e_1 \otimes f_2, \dots, e_2 \otimes f_m,\ \dots,\ e_n \otimes f_1,\ e_1 \otimes f_2, \dots, e_n \otimes f_m,$$

etc. for $\{e'_i \otimes f'_k\}$. Thus, $e_j \otimes f_l$ is the $\big((j-1)m + l\big)^{\text{th}}$ element of our chosen ordered basis of $V \otimes W$, where $1 \le j \le n$ and $1 \le l \le m$. Similarly, $e'_i \otimes f'_k$ is the $\big((i-1)m' + k\big)^{\text{th}}$ element of our chosen ordered basis of $V' \otimes W'$, where $1 \le i \le n'$ and $1 \le k \le m'$. Thus the number $\tilde{A}^i{}_j \tilde{B}^k{}_l$ in equation (2.20) is the entry in row $(i-1)m' + k$ and column $(j-1)m + l$ of the $(n'm') \times (nm)$ matrix of $A \otimes B$ with respect to the ordered bases above.[5] Letting $\tilde{A} \otimes \tilde{B}$ denote this matrix, we can write $\tilde{A} \otimes \tilde{B}$ in block form as

$$\begin{bmatrix} \tilde{A}^1{}_1 \tilde{B} & \tilde{A}^1{}_2 \tilde{B} & \dots & \tilde{A}^1{}_n \tilde{B} \\ \tilde{A}^2{}_1 \tilde{B} & \tilde{A}^2{}_2 \tilde{B} & \dots & \tilde{A}^2{}_n \tilde{B} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{A}^{n'}{}_1 \tilde{B} & \tilde{A}^{n'}{}_2 \tilde{B} & \dots & \tilde{A}^{n'}{}_n \tilde{B} \end{bmatrix}. \tag{2.21}$$

Note that, up to our choice to order a tensor-product basis lexicographically, (2.21) followed directly from the standard convention for defining the matrix of a linear transformation with respect to bases of the domain and codomain. In view of this fact, given any matrices $\tilde{A} \in M_{n' \times n}(\mathbf{R}), \tilde{B} \in M_{m' \times m}(\mathbf{R})$, we *define* the tensor product $\tilde{A} \otimes \tilde{B}$ to be the matrix (2.21) (even absent any explicit reference to linear transformations). ▲

Remark 2.32 As a practical matter, the most important tensor products are those of finite-dimensional vector spaces. In a first introduction to tensor products in the finite-dimensional setting, given finite-dimensional vector spaces V and W, the space "$V \otimes W$" is usually not defined directly. Instead the space that is defined directly is $V^* \otimes W^*$, which is taken to be the space $\mathrm{Bihom}(V \times W, \mathbf{R})$, a space to which $V^* \otimes W^*$ is

[5] LaTeX note: Ordinarily, to prevent LaTeX from undesirably stacking a first-index superscript directly over a second-index subscript (which would eliminate the distinction between which index is first and which is second), protecting the superscripted expression with curly braces suffices. For example, you can produce "$A^i{}_j$" with " `$ { A^i }_ j $`." However, when there is a tilde over the A, LaTeX becomes very insistent on stacking the superscript directly over the subscript. To produce the output "$\tilde{A}^i{}_j$", I used " `$ { A^i}_{\ j} $`."

canonically isomorphic (see Proposition 2.30). In such a first introduction, the space $V \otimes W$ is usually not of direct importance, but, by the preceding, can be taken to be the space $\operatorname{Bihom}(V^* \times W^*, \mathbf{R})$, since $V^{**} \underset{\text{canon.}}{\cong} V$ and $W^{**} \underset{\text{canon.}}{\cong} W$. This "quick-and-dirty" approach allows much faster derivations of many results in these notes; it really cannot be matched in efficiency.

But despite its efficiency, the quick-and-dirty approach is a bit of a cheat—even though finite-dimensional vector spaces are the most important setting for tensor products—and has several conceptual drawbacks:

- The notation "$V^* \otimes W^*$" suggests that "$\otimes$" (in this usage) is a binary operation, producing a new vector space from the spaces V^* and W^*. But the individual vector spaces V^* and W^* do not even enter the definition of $\operatorname{Bihom}(V \times W, \mathbf{R})$, leaving one with the nagging question of how to view "$\otimes$" as an operation on V^* and W^*.

- It feels unsatisfying that the definition of $V \otimes W$ should rely on the duals of these spaces.

The fact that Definition 2.1 leads to the universal property in Proposition 2.10—capturing the essence of bilinearity, a concept that does not depend on finite-dimensionality—shows that Definition 2.1 should be viewed as the "true" or "correct" definition of tensor product. With this understanding, observe that showing the equivalence of the "quick-and-dirty" definition of $V \otimes W$ to Definition 2.1 requires *two* uses of finite-dimensionality: one to show that the canonical injection $V^* \otimes W^* \hookrightarrow \operatorname{Bihom}(V \times W, \mathbf{R})$ is an isomorphism, and one to show that $V^{**} \underset{\text{canon.}}{\cong} V$ and $W^{**} \underset{\text{canon.}}{\cong} W$. ▲

3 Tensor product of more than two spaces

3.1 Definitions and the universal property

Tensor products of more than two vector spaces are defined analogously to the two-space case, and have analogous properties. We will run through the basics quickly and informally.

Let $V_1, \ldots, V_k$ be vector spaces. Let $\mathcal{I}(V_1, \ldots, V_k) \subset \mathbf{R}[V_1 \times \cdots \times V_k]$ be the subspace generated by all elements of any of the following forms (written in "formal linear combination" notation):

$$
\begin{aligned}
\{(v_1,\ldots,v_{i-1},w+w',v_{i+1},\ldots,v_k) &- (v_1,\ldots,v_{i-1},w,v_{i+1},\ldots,v_k)\\
&- (v_1,\ldots,v_{i-1},w',v_{i+1},\ldots,v_k),\\
&\quad \text{where } i\in\{1,\ldots,k\},\ w,w'\in V_i,\\
&\quad \text{and } v_j\in V_j \text{ for } j\in\{1,\ldots,k\}\setminus\{j\},\\
\{(v_1,\ldots,v_{i-1},cv_i,v_{i+1},\ldots,v_k) &- c(v_1,\ldots,v_{i-1},v_i,v_{i+1},\ldots,v_k),\\
&\quad \text{where } i\in\{1,\ldots,k\},\\
&\quad v_j\in V_j \text{ for } j\in\{1,\ldots,k\},\\
&\quad \text{and } c\in\mathbf{R}.
\end{aligned}
$$

(Above, it is understood that if $i = 1$ [respectively $i = k$], then the vectors v_{i-1} [resp. v_{i+1}] are omitted.) We then define

$$
\begin{aligned}
V_1\otimes\ldots\otimes V_k := \otimes^{(k)}(V_1,\ldots,V_k) &:= \mathbf{R}[V_1\times\cdots\times V_k]/\mathcal{I}(V_1,\ldots,V_k);\\
\pi:\mathbf{R}[V_1\times\cdots\times V_k] &\to V_1\otimes\ldots\otimes V_k\,,\quad \text{the quotient map;}\\
\iota: V_1\times\ldots\times V_k &\to \mathbf{R}[V_1\times\cdots\times V_k]\,,\quad \text{the natural inclusion;}\\
v_1\otimes\ldots\otimes v_k &:= \pi(\iota(v_1,\ldots,v_k)) := \otimes^{(k)}_{\mathrm{op}}(v_1,\ldots,v_k)\\
&\qquad \text{for } v_1\in V_1,\ldots,v_k\in V_k\,.
\end{aligned}
$$

Thus $\pi\left(\sum_\mu c_\mu(v_1^{(\mu)},\ldots,v_k^{(\mu)})\right) = \sum_\mu c_\mu v_1^{(\mu)}\otimes\ldots\otimes v_k^{(\mu)}$.

Remark 3.1 Note that $\otimes^{(2)}(V_1,V_2) = V_1\otimes V_2$ and that $\otimes^{(2)}_{\mathrm{op}} : V_1\times V_2\to V_1\otimes V_2$ is the map called simply $\otimes_{\mathrm{op}}$ in Section 2. We can even take $k = 1$ above: $\otimes^{(1)}(V) = V$ and $\otimes^{(1)}_{\mathrm{op}} : V \to V$ is simply the identity map id_V. But note that an expression such as $v_1\otimes v_2\otimes\ldots\otimes v_k$ is, so far, an "inseparable word"; for example, the "sub-word" $v_1\otimes v_2$ does not yet mean $\otimes^{(2)}_{\mathrm{op}}(v_1,v_2)$. Interpreting such sub-words the way the notation suggests assumes that, in the notation $v_1\otimes v_2\otimes\ldots\otimes v_k$, the symbol "$\otimes$" has an associativity property. The notation $v_1\otimes v_2\otimes\ldots\otimes v_k$ has been chosen to reflect associativity that we will *eventually* see that the symbol $\otimes$ enjoys, but that we have not addressed yet. We address associativity in Section 3.2.

Corollary 3.2 *The map $\otimes^{(k)}_{\mathrm{op}} : V_1\times\cdots\times V_k \to V_1\otimes\ldots\otimes V_k$ is multilinear* (i.e. linear in each variable with the others held fixed)*, and its image spans $V_1\otimes\ldots\otimes V_k$.*

Proof: Exercise. ■

Note that for $k = 1$, "multilinear" simply means "linear".

Proposition 3.3 (Universal Property of Multiple Tensor Products) *Let $V_1,\ldots,V_k$ be vector spaces.*

(a) The triple $(V_1 \times \cdots \times V_k,\ V_1 \otimes \ldots \otimes V_k,\ \otimes_{\rm op}^{(k)})$ *has the following universal property: for any vector space* Z*, and any* *<u>multi</u>linear map* $B : V_1 \times \cdots \times V_k \to Z$*, there exists a unique* *<u>linear</u> map* $L_B : V_1 \otimes \ldots \otimes V_k \to Z$ *such that* $L_B(v_1 \otimes \ldots \otimes v_k) = B(v_1, \ldots, v_k)$ *for all* $v_1 \in V_1, \ldots, v_k \in V_k$ (equivalently, such that $B = L_B \circ \otimes_{\rm op}^{(k)}$, as indicated by the commutative diagram in Figure 2)*.*

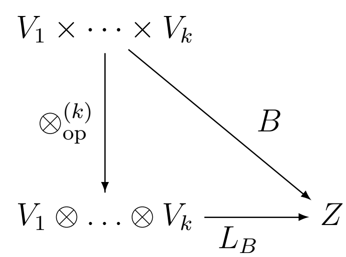

Figure 2: Diagram for Proposition 3.3(a)

(b) The pair $(\otimes_{\rm op}^{(k)}, V_1 \otimes \ldots \otimes V_k)$ *is "unique up to isomorphism", in the following sense: if* X *is a vector space and* $T : V_1 \times \cdots \times V_k \to X$ *is another multilinear map such that* $(V_1 \otimes \ldots \otimes V_k, X, T)$ *has the universal property described in part (a), then there is an isomorphism* $L : V_1 \otimes \ldots \otimes V_k \to X$ *such that* $T = L \circ \otimes_{\rm op}^{(k)}$*.*

Proof: Exercise. ■

Corollary 3.4 *Let* $V_1, \ldots V_k$ *and* $V_1', \ldots V_k'$ *be vector spaces and let* $A_i : V_i \to V_i'$ *be linear maps,* $1 \leq i \leq k$*. Then there is a unique linear map* $L : V_1 \otimes \ldots \otimes V_k \to V_1' \otimes \ldots \otimes V_k'$ *satisfying* $L(v_1 \otimes \ldots \otimes v_k) = (A_1 v_1) \otimes \ldots \otimes (A_k v_k)$ *for all* $(v_1, \ldots, v_k) \in V_1 \times \cdots \times V_k$*.*

Proof: Exercise. ■

Exercise 3.5 Let $V_1, \ldots, V_k$ be vector spaces. It is easily seen that $\mathrm{Multihom}(V_1 \times \cdots \times V_k, \mathbf{R}) := k\text{-}\mathrm{Hom}(V_1 \times \cdots \times V_k, \mathbf{R}) := \{$multilinear maps $V_1 \times \cdots \times V_k \to \mathbf{R}\}$ is a subspace of $\mathbf{R}^{V_1 \times \cdots \times V_k}$.

Generalize Proposition 2.25abc:

(a) Exhibit, with proof, a canonical isomorphism

$$\iota : \mathrm{Multihom}(V_1 \times \cdots \times V_k, \mathbf{R}) \to (V_1 \otimes \ldots \otimes V_k)^* \,.$$

(b) Exhibit, with proof, a generalization of the linear map j in Proposition 2.25(b) to an injective linear map $j : V_1^* \otimes \ldots \otimes V_k^* \to \mathrm{Multihom}(V_1 \times \cdots \times V_k, \mathbf{R})$.

(c) Conclude that $j \circ \iota : V_1^* \otimes \ldots \otimes V_k^* \to (V_1 \otimes \ldots \otimes V_k)^*$ is a canonical linear injection.

Definition 3.6 (tensor product of k linear maps) Notation as in Corollary 3.4. The linear map L is called the *tensor product* of the linear maps $A_1, \ldots A_k$, and is denoted $A_1 \otimes \ldots \otimes A_k$.

Thus, $A_1 \otimes \ldots \otimes A_k : V_1 \otimes \ldots \otimes V_k \to V_1' \otimes \ldots \otimes V_k'$ is the unique linear map satisfying

$$(A_1 \otimes \ldots \otimes A_k)(v_1 \otimes \ldots \otimes v_k) = A_1 v_1 \otimes \ldots \otimes A_k v_k \tag{3.22}$$

for all $(v_1, \ldots, v_k) \in V_1 \times \cdots \times V_k$.

Proposition 3.7 *Let $V_1, \ldots, V_k$ be vector spaces and, for each i, let $\mathcal{B}_i$ be a basis of V_i. Then the set*

$$\{v_1 \otimes \ldots \otimes v_k \mid v_i \in \mathcal{B}_i\,,\ 1 \le i \le k\} \tag{3.23}$$

is a basis of $V_1 \otimes \ldots \otimes V_k$.

Proof: Exercise. (Extend proofs of Proposition 2.23 and Corollary 2.24 using induction.) ■

Exercise 3.8 Notation as in Exercise 3.5, but assume now that $V_1, \ldots, V_k$ are *finite-dimensional*. Show that the linear map j in Exercise 3.5 is an isomorphism, and hence that the three spaces $V_1^* \otimes \ldots \otimes V_k^*$, Multihom$(V_1 \times \cdots \times V_k, \mathbf{R})$, and $(V_1 \otimes \ldots \otimes V_k)^*$ are canonically isomorphic to each other.

Exercise 3.9 Let the category $\mathcal{C}'$ be as defined in Exercise 2.21, and let $(\mathcal{C}')^k$ be the product category $\underbrace{\mathcal{C}' \times \mathcal{C}' \times \cdots \times \mathcal{C}'}_{k \text{ copies}}$. Generalize the result of Exercise 2.21(b): show that the relationships among k-fold tensor product of vector spaces, linear transformations, and elements, can be encoded by a covariant functor $\bigotimes^{(k)} : (\mathcal{C}')^k \to \mathcal{C}'$.

3.2 Associativity

Proposition 3.10 *Let V_1, V_2, V_3 be vector spaces. Then the natural bijections $(V_1 \times V_2) \times V_3 \leftrightarrow V_1 \times V_2 \times V_3 \leftrightarrow V_1 \times (V_2 \times V_3)$ induce isomorphisms*

$$\begin{aligned} \iota_1 : V_1 \otimes V_2 \otimes V_3 &\to (V_1 \otimes V_2) \otimes V_3, \\ \iota_2 : V_1 \otimes V_2 \otimes V_3 &\to V_1 \otimes (V_2 \otimes V_3) \end{aligned}$$

satisfying

$$\begin{aligned} \iota_1(v_1 \otimes v_2 \otimes v_3) &= (v_1 \otimes v_2) \otimes v_3, \\ \iota_2(v_1 \otimes v_2 \otimes v_3) &= v_1 \otimes (v_2 \otimes v_3), \end{aligned}$$

for all $(v_1, v_2, v_3) \in V_2 \times V_2 \times V_3$.

In particular, the three spaces $V_1 \otimes V_2 \otimes V_3$, $(V_1 \otimes V_2) \otimes V_3$, and $V_1 \otimes (V_2 \otimes V_3)$ are canonically isomorphic.

Proof: The composition

$$V_1 \times V_2 \times V_3 \underset{\substack{\text{natural}\\ \text{bijection}}}{\longrightarrow} (V_1 \times V_2) \times V_3 \underset{\otimes_{\text{op}}^{(2)} \times \text{id}_{V_3}}{\longrightarrow} (V_1 \otimes V_2) \times V_3 \underset{\otimes_{\text{op}}^{(2)}}{\longrightarrow} (V_1 \otimes V_2) \otimes V_3$$

is simply the trilinear map $(v_1, v_2, v_3) \mapsto (v_1 \otimes v_2) \otimes v_3$. By Proposition 3.3, this trilinear map determines a unique linear map $\iota_1 : V_1 \otimes V_2 \otimes V_3 \to (V_1 \otimes V_2) \otimes V_3$ for which $\iota_1(v_1 \otimes v_2 \otimes v_3) = (v_1 \otimes v_2) \otimes v_3$. Similarly, there we obtain a canonical linear map $\iota_2 : V_1 \times V_2 \times V_3 \to V_1 \otimes (V_2 \otimes V_3)$.

Let $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ be bases for V_1, V_2, V_3 respectively. By Corollary 2.24 (or Proposition 3.7 with $k = 2$), the set $\{b_1 \otimes b_2 \mid (b_1, b_2) \in \mathcal{B}_1 \times \mathcal{B}_2\}$ is a basis of $V_1 \otimes V_2$, and, consequently, $\mathcal{B} := \{(b_1 \otimes b_2) \otimes b_3 \mid (b_1, b_2, b_3) \in \mathcal{B}_1 \times \mathcal{B}_2 \times \mathcal{B}_3\}$ is a basis of $(V_1 \otimes V_2) \otimes V_3$. By Proposition 3.7, $\mathcal{B}' := \{\otimes_{\text{op}}^{(3)}(b_1, b_2, b_3) = b_1 \otimes b_2 \otimes b_3 \mid (b_1, b_2, b_3) \in \mathcal{B}_1 \times \mathcal{B}_2 \times \mathcal{B}_3\}$ is a basis of $V_1 \otimes V_2 \otimes V_3$. Since ι_1 carries the basis $\mathcal{B}$ bijectively to the basis $\mathcal{B}'$, the map ι_1 is an isomorphism. Similarly, ι_2 is an isomorphism. ■

It is easily seen that Proposition 3.10 generalizes. For example, for any $k, l \in \mathbf{N}$ with $k < l$, and vector spaces $V_1, \ldots, V_l$ the natural bijection

$$(V_1 \times \cdots \times V_k) \times (V_{k+1} \times \cdots \times V_l) \to V_1 \times \cdots \times V_k \times V_{k+1} \times \cdots \times V_l$$

induces a canonical isomorphism

$$(V_1 \otimes \ldots \otimes V_k) \otimes (V_{k+1} \otimes \ldots \otimes V_l) \underset{\cong}{\to} V_1 \otimes \ldots \otimes V_l \tag{3.24}$$

satisfying

$$\begin{aligned} (v_1 \otimes \ldots \otimes v_k) \otimes (v_{k+1} \otimes \ldots \otimes v_l) &\mapsto v_1 \otimes \ldots \otimes v_l \\ \text{(equivalently,} \quad \otimes_{\text{op}}^{(2)} \left(\otimes_{\text{op}}^{(k)} (v_1, \ldots, v_k), \otimes_{\text{op}}^{(l)}(v_{k+1}, \ldots, v_l)\right) &= \otimes_{\text{op}}^{(k+l)}(v_1, \ldots, v_{k+l}) \ ) \end{aligned} \tag{3.25}$$

for all $(v_1, \ldots, v_l) \in V_1 \times \cdots \times V_l$.

Convention 3.11 Since the canonical isomorphisms in Proposition 3.10 and, more generally, in (3.24), are induced simply by re-parenthesizing Cartesian products of the vector spaces V_i, we will regard them as *equalities* (cf. the third paragraph of Remark 2.18); e.g. we will write

$$(V_1 \otimes \ldots \otimes V_k) \otimes (V_{k+1} \otimes \ldots \otimes V_l) = V_1 \otimes \ldots \otimes V_l \tag{3.26}$$

and, for elements $v_i \in V_i$, write

$$(v_1 \otimes \ldots \otimes v_k) \otimes (v_{k+1} \otimes \ldots \otimes v_l) = v_1 \otimes \ldots \otimes v_l \ . \tag{3.27}$$

(This abuse of notation can be avoided by introducing appropriate equivalence relations on appropriate sets, and defining tensor product of equivalence classes. We are choosing not to add that baggage to this presentation. This spares us from having to penetrate the potential forest of notation illustrated in (3.25).)

Remark 3.12 ...[Why I didn't just define tensor product of more than two vector spaces by iterating the binary operations $(V, W) \to V \otimes W$ and $(v, w) \mapsto v \otimes w$.]

Remark 3.13 ...[The appropriate categories and functors.]

4 Tensor product of multiple copies of the same vector space

Throughout this section:

- V is a vector space.
- $k \in \mathbf{N}$.
- $S_k :=$ {permutations of $\{1, \dots, k\}$ }, the *symmetric group on k slements.*
- The sign of a permutation $\sigma \in S_k$ is denoted $(-1)^\sigma$.
- V^k denotes the k-fold Cartesian product $\underbrace{V \times \cdots \times V}_{k \text{ copies}}$
- $V^{\otimes k}$ denotes the k-fold tensor product $\underbrace{V \otimes \ldots \otimes V}_{k \text{ copies}}$.

4.1 Action of the symmetric group

The symmetric group S_k acts on V^k in a natural way: for $\sigma \in S_k$ and $v := (v_1, \dots, v_k) \in V^k$, we define

$$\tilde{T}_\sigma(v) := \sigma \cdot v := \sigma \cdot (v_1, \dots, v_k) := (v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(k)}); \tag{4.28}$$

i.e. $(\sigma \cdot v)_i = v_{\sigma^{-1}(i)}$. With this definition—which is the one for which the action of σ on v moves v_i to the $\sigma(i)^{\text{th}}$ slot—we have $\tilde{T}_{\rho\sigma} = \tilde{T}_\rho \circ \tilde{T}_\sigma$; i.e. the action by S_k is a left-action, consistent with our notation ("$\sigma \cdot v$" rather than "$v \cdot \sigma$").[6] Note that v_i is the i^{th} *vector* in a k-tuple v of vectors; there are no *components* of vectors, relative to whatever basis, involved in these formulas.

For each $\sigma \in S_k$, the map

$$\begin{aligned} \otimes^{(k)}_{\text{op}} \circ \tilde{T}_\sigma : V^k &\to V^{\otimes k}, \\ (v_1, \dots, v_k) &\mapsto v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(k)}, \end{aligned}$$

[6]For purposes of these notes, it is immaterial whether we use the natural left action or natural right action, the latter being defined by $(v \cdot \sigma)_i = v_{\sigma(i)}$. define $(\sigma \cdot v)_i$ to be $v_{\sigma^{-1}(i)}$ or $v_{\sigma(i)}$. Our choice to use the left-action was simply a matter of notational preference.

is clearly multilinear, and hence determines a unique linear map $T_\sigma : V^{\otimes k} \to V^{\otimes k}$ satisfying $T_\sigma \circ \otimes_{\text{op}}^{(k)} = \otimes_{\text{op}}^{(k)} \circ \tilde{T}_\sigma$. For $\rho, \sigma \in S_k$ we have

$$T_\rho \circ T_\sigma \circ \otimes_{\text{op}}^{(k)} = T_\rho \circ \otimes_{\text{op}}^{(k)} \circ \tilde{T}_\sigma = \otimes_{\text{op}}^{(k)} \circ \tilde{T}_\rho \circ \tilde{T}_\sigma = \otimes_{\text{op}}^{(k)} \circ \tilde{T}_{\rho\sigma} = T_{\rho\sigma} \circ \otimes_{\text{op}}^{(k)} \tag{4.29}$$

Since the image of $\otimes_{\text{op}}^{(k)}$ spans $V^{\otimes k}$, (4.29) shows that $T_{\rho\sigma} = T_\rho \circ T_\sigma$, i.e. that the map $S_k \times V^k \to V^k$ given by $(\sigma, h) \mapsto T_\sigma(h)$ is a left-action of S_k on $V^{\otimes k}$. For this reason we will generally use the notation "$\sigma \cdot$" for $T_\sigma : V^{\otimes k} \to V^{\otimes k}$, as well as for $\tilde{T}_\sigma : V^k \to V^k$.

Definition 4.1 Let $v_1, \ldots, v_k \in V$.

1. The *symmetric product* of $v_1, \ldots, v_k$ is

$$v_1 \odot v_2 \odot \cdots \odot v_k := \sum_{\sigma \in S_k} \sigma \cdot (v_1 \otimes \ldots \otimes v_k) = \left(\sum_{\sigma \in S_k} T_\sigma \right) (v_1 \otimes \ldots \otimes v_k) \quad \in V^{\otimes k}. \tag{4.30}$$

2. The *wedge product* of $v_1, \ldots, v_k$ (in the indicated order) is

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k := \sum_{\sigma \in S_k} (-1)^\sigma \sigma \cdot (v_1 \otimes \ldots \otimes v_k) = \left(\sum_{\sigma \in S_k} (-1)^\sigma T_\sigma \right) (v_1 \otimes \ldots \otimes v_k) \quad \in V^{\otimes k}. \tag{4.31}$$

▲

For example,

$$\begin{aligned} v_1 \odot v_2 &= v_1 \otimes v_2 + v_2 \otimes v_1, \\ v_1 \wedge v_2 &= v_1 \otimes v_2 - v_2 \otimes v_1, \end{aligned}$$

and

$$\begin{aligned} v_1 \wedge v_2 \wedge v_3 = \; & v_1 \otimes v_2 \otimes v_3 + v_2 \otimes v_3 \otimes v_1 + v_3 \otimes v_1 \otimes v_2 \\ & - v_1 \otimes v_3 \otimes v_2 - v_2 \otimes v_1 \otimes v_3 - v_3 \otimes v_2 \otimes v_1 \ . \end{aligned}$$

Remark 4.2 Some people prefer to normalize the definition of $v_1 \odot v_2 \odot \cdots \odot v_k$ by putting a "$\frac{1}{k!}$" in front of the sum over the symmetric group, so that for the k-fold symmetric product of a vector with itself, one has $v \odot v \odot \cdots \odot v = v \otimes v \otimes \ldots \otimes v$.

Exercise 4.3 Let $v_1, \ldots, v_k \in V$ and let $\sigma \in S_k$. Show that

$$\sigma \cdot (v_1 \odot v_2 \odot \cdots \odot v_k) = v_1 \odot v_2 \odot \cdots \odot v_k$$

and that

$$\sigma \cdot (v_1 \wedge v_2 \wedge \cdots \wedge v_k) = (-1)^\sigma v_1 \wedge v_2 \wedge \cdots \wedge v_k \ .$$

4.2 Symmetric powers $\mathrm{Sym}^k(V)$ and their universal property

The k^{th} *symmetric power* of V, or k-fold *symmetric tensor product* of V, is the set

$$\mathrm{Sym}^k(V) := \{h \in V^{\otimes k} : \sigma \cdot h = h \text{ for all } \sigma \in S_k\}. \tag{4.32}$$

Exercise 4.4 (a) Show that $\mathrm{Sym}^k(V)$ is a subspace of $V^{\otimes k}$.

(b) From Definition 4.1, it is clear that the map $V^k \to V^{\otimes k}$ given by $(v_1, \dots, v_k) \mapsto v_1 \odot \cdots \odot v_k$ is multilinear, and hence determines a linear map $P_k : V^{\otimes k} \to V^{\otimes k}$. Show that the image of P_k is precisely $\mathrm{Sym}^k(V)$, and that if $h \in \mathrm{Sym}^k(V)$, then $P_k(h) = c(k)h$, where $c(k) = k!$. (Thus, up to a constant factor, P_k is a projection map from $V^{\otimes k}$ onto the subspace $\mathrm{Sym}^k(V)$.)

Definition 4.5 A multilinear map $B : V^k \to Z$ (where Z is a vector space) is *symmetric* if $B(\sigma \cdot v) = B(v)$ for all $\sigma \in S_k$ and all $v \in V^k$.

Note that since transpositions generate the symmetric group, a sufficient condition for B to be symmetric is that $B(\sigma \cdot v) = B(v)$ for all *transpositions* in S_k and all $v \in V^k$.

Proposition 4.6 (Universal Property of Symmetric Tensor Product) *Let $\odot^{(k)} : V^k \to \mathrm{Sym}^k(V)$ be the map $(v_1, \dots, v_k) \mapsto v_1 \odot \cdots \odot v_k$.*

(a) The triple $(V^k, \mathrm{Sym}^k(V), \odot^{(k)})$ has the following universal property: for any vector space Z, and any symmetric *multilinear map $B : V^k \to Z$, there exists a unique linear map $L_B : \mathrm{Sym}^k(V) \to Z$ such that $L_B(v_1 \odot \cdots \odot v_k) = B(v_1, \dots, v_k)$ for all $(v_1, \dots, v_k) \in V^k$* (equivalently, such that $B = L_B \circ \odot^{(k)}$, as indicated by the commutative diagram in Figure 3).

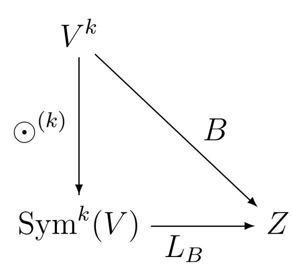

Figure 3: Diagram for Proposition 4.6(a)

(b) The pair $(\odot^{(k)}, \mathrm{Sym}^k(V))$ is "unique up to isomorphism", in the following sense: if X is a vector space and $T : V^k \to X$ is another symmetric multilinear map such that (V^k, X, T) has the universal property described in part (a), then there is an isomorphism $L : \mathrm{Sym}^k(V) \to X$ such that $T = L \circ \odot^{(k)}$.

Proof: Exercise. ■

Corollary 4.7 *Let V, W be vector spaces and let $A : V \to W$ be a linear map. Then there is a unique linear map $L : \mathrm{Sym}^k(V) \to \mathrm{Sym}^k(W)$ satisfying $L(v_1 \odot \cdots \odot v_k) = Av_1 \odot \cdots \odot Av_k$ for all $(v_1, \ldots, v_k) \in V^k$.*

Proof: Exercise. ■

Definition 4.8 (symmetric powers of a linear map) Notation as in Corollary 4.7. We call the linear map L the k^{th} *symmetric power* of A, and denote it $\mathrm{Sym}^k(A)$.

Thus, $\mathrm{Sym}^k(A) : \mathrm{Sym}^k(V) \to \mathrm{Sym}^k(W)$ is the unique linear map satisfying

$$(\mathrm{Sym}^k(A))(v_1 \odot \cdots \odot v_k) = Av_1 \odot \cdots \odot Av_k \tag{4.33}$$

for all $(v_1, \ldots, v_k) \in V^k$.

Proposition 4.9 *Assume that V has finite dimension n, and let $\{e_i\}_{i=1}^n$ be a basis of V. Then the set*

$$\{e_{i_1} \odot \cdots \odot e_{i_k} \mid 1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n\} \tag{4.34}$$

is a basis of $\mathrm{Sym}^k(V)$, *and* $\dim(\mathrm{Sym}^k(V)) = \binom{n+k-1}{k}$.

Proof: Exercise. ■

Exercise 4.10 Show that if V is finite-dimensional, then $\mathrm{Sym}^k(V^*)$ is canonically isomorphic to $(\mathrm{Sym}^k(V))^*$.

4.3 Exterior powers $\bigwedge^k(V)$ and their universal property

The k^{th} *exterior power of* V, or k-fold *wedge product* of V, is the set

$$\textstyle\bigwedge^k(V) := \{\xi \in V^{\otimes k} : \sigma \cdot \omega = (-1)^\sigma \xi \text{ for all } \sigma \in S_k\}. \tag{4.35}$$

Exercise 4.11 (a) Show that $\bigwedge^k(V)$ is a subspace of $V^{\otimes k}$.

(b) From Definition 4.1, it is clear that the map $V^k \to V^{\otimes k}$ given by $(v_1, \ldots, v_k) \mapsto v_1 \wedge \cdots \wedge v_k$ is multilinear, and hence determines a linear map $\mathrm{Alt}_k : V^{\otimes k} \to V^{\otimes k}$. Show that the image of Alt_k is precisely $\bigwedge^k(V)$, and that if $\xi \in \bigwedge^k(V)$, then $\mathrm{Alt}_k(\xi) = c(k)\xi$, where $c(k) = k!$ (Thus, up to a constant factor, Alt_k is a projection map from $V^{\otimes k}$ onto the subspace $\bigwedge^k(V)$.)

Definition 4.12 A multilinear map $B : V^k \to Z$ (where Z is a vector space) is called *alternating*, or *antisymmetric*, or *totally antisymmetric*, if $B(\sigma \cdot v) = (-1)^\sigma B(v)$ for all $\sigma \in S_k$ and all $v \in V^k$.

Note that since transpositions generate the symmetric group, and the sign-homomorphism $\sigma \mapsto (-1)^\sigma$ is (indeed) a homomorphism $S_k \to \{\pm 1, \cdot\}$, a sufficient condition for B to be symmetric is that $B(\sigma \cdot v) = -B(v)$ for all *transpositions* in S_k and all $v \in V^k$.

Proposition 4.13 (Universal Property of Wedge Product)
Let $\wedge^{(k)} : V^k \to \bigwedge^k(V)$ be the map $(v_1, \ldots, v_k) \mapsto v_1 \wedge \cdots \wedge v_k$.

(a) The triple $(V^k, \bigwedge^k(V), \wedge^{(k)})$ has the following universal property: for any vector space Z, and any alternating *multilinear map $B : V^k \to Z$, there exists a unique linear map $L_B : \bigwedge^k(V) \to Z$ such that $L_B(v_1 \wedge \cdots \wedge v_k) = B(v_1, \ldots, v_k)$ for all* $(v_1, \ldots, v_k) \in V^k$ (equivalently, such that $B = L_B \circ \wedge^{(k)}$, as indicated by the commutative diagram in Figure 4).

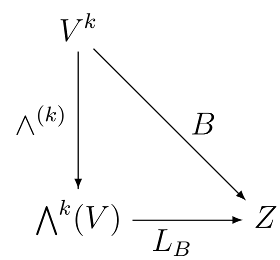

Figure 4: Diagram for Proposition 4.13(a)

(b) The pair $(\wedge^{(k)}, \bigwedge^k(V))$ is "unique up to isomorphism", in the following sense: if X is a vector space and $T : V^k \to X$ is another symmetric multilinear map such that (V^k, X, T) has the universal property described in part (a), then there is an isomorphism $L : \bigwedge^k(V) \to X$ such that $T = L \circ \wedge^{(k)}$.

Proof: Exercise. ■

Corollary 4.14 *Let V, W be vector spaces and let $A : V \to W$ be a linear map. Then there is a unique linear map $L : \bigwedge^k(V) \to \bigwedge^k(W)$ satisfying $L(v_1 \wedge \cdots \wedge v_k) = Av_1 \wedge \cdots \wedge Av_k$ for all $(v_1, \ldots, v_k) \in V^k$,*

Proof: Exercise. ■

Exercise 4.15 (a) Show that for $k > 1$, $\mathrm{Sym}^k(V) \cap \bigwedge^k(V) = \{0\}$.

(b) Show that $V \otimes V = \mathrm{Sym}^2(V) \oplus \bigwedge^2(V)$ (direct sum of subspaces).

(c) Show that part (b) does not extend to higher-order tensor products: if $k > 2$ and $\dim(V) > 1$, then $V^{\otimes k}$ is not spanned by the subspaces $\mathrm{Sym}^k(V)$ and $\bigwedge^k(V)$.

Definition 4.16 (exterior powers of a linear map) Notation as in Corollary 4.14. We call the linear map L the k^{th} *exterior power* of A, and denote it $\bigwedge^k(A)$.

Thus, $\bigwedge^k(A) : \bigwedge^k(V) \to \bigwedge^k(W)$ is the unique linear map satisfying

$$(\textstyle\bigwedge^k(A))(v_1 \wedge \cdots \wedge v_k) = Av_1 \wedge \cdots \wedge Av_k \tag{4.36}$$

for all $(v_1, \ldots, v_k) \in V^k$.

Proposition 4.17 *Assume that V has finite dimension n, and let $\{e_i\}_{i=1}^n$ be a basis of V. For $k \le n$, the set*

$$\{e_{i_1} \wedge \cdots \wedge e_{i_k} \mid 1 \le i_1 < i_2 < \cdots < i_k \le n\} \tag{4.37}$$

is a basis of $\bigwedge^k(V)$, and $\dim(\bigwedge^k(V)) = \binom{n}{k}$.

Proof: Exercise. ■

Exercise 4.18 Let $k \in \mathbf{N}$.

(a) Show that if V is finite-dimensional, then $\bigwedge^k(V^*)$ is canonically isomorphic to $(\bigwedge^k(V))^*$, and also canonically isomorphic to the space

$$\mathcal{A}^k(V) := \{\omega : V^k \to \mathbf{R} \mid \omega \text{ is multilinear and alternating}\}.$$

(The notation is "local" to these notes, not standard.) See Remark 4.19.

(b) Define a map

$$\begin{array}{rcl} \tilde{\wedge}^{(k)} : (V^*)^k & \to & \mathcal{A}^k(V), \\ (\eta_1, \ldots, \eta_k) & \mapsto & \eta_1 \,\tilde{\wedge}\, \eta_2 \,\tilde{\wedge} \ldots \tilde{\wedge}\, \eta_k \ , \end{array}$$

by

$$\eta_1 \,\tilde{\wedge}\, \eta_2 \,\tilde{\wedge} \ldots \tilde{\wedge}\, \eta_k \,(v_1, v_2, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^\sigma \, \eta_1(v_{\sigma_1})\, \eta_2(v_{\sigma_2}) \ldots \eta_k(v_{\sigma(k)})$$

Let $\iota : \bigwedge^k(V^*) \to \mathcal{A}^k(V)$ be the canonical isomorphism you used to prove part (a). Show that the following diagram (Figure 5) commutes:

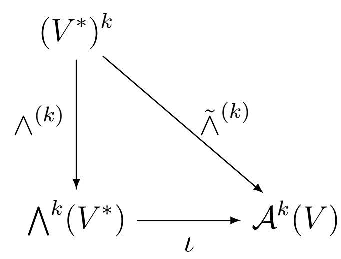

Figure 5: Diagram for Exercise 4.18(b)

Remark 4.19 In the introduction to exterior algebra usually presented in an introduction to differential forms on manifolds, if M is manifold and $p \in M$, the space $\bigwedge^k(T_p^*M)$ is usually defined to *be* the space $\mathcal{A}^k(T_pM)$.

Exercise 4.20 The purpose of this exercise is to illustrate how much simpler the universal properties in Propositions 2.10, 3.3, 4.6, and 4.13 can make certain arguments. For simplicity, we illustrate this only for the universal property of exterior powers (and only for finite-dimensional vector spaces). In this exercise, the goal is to prove Corollary 4.14 for finite-dimensional V and W **without making use any of the universal properties mentioned above.**

To make things as simple as possible, assume for this exercise that, for a finite-dimensional vector space V, the second result in Exercise 4.18(a) is taken as the *definition* of $\bigwedge^k(V^*)$ (cf. Remark 4.19): $\bigwedge^k(V^*) := \mathcal{A}^k(V)$. In view of this definition and the conclusion of Exercise 4.18(b), replace the notation "$\tilde{\wedge}$" in Exercise 4.18(b) by "$\wedge$".

Note that it suffices to prove Corollary 4.14 with the general finite-dimensional vector spaces V and W replaced by their duals, since $V^{**} := (V^*)^*$ is canonically isomorphic to V in the finite-dimensional case.

(a) Let V be vector space with finite, positive dimension n, let $\mathcal{B} := \{\theta^i\}_{i=1}^n$ be a basis of V^*, and assume $k \le n$. **Without using Proposition 4.17**, prove that the set

$$\{\theta^{i_1} \wedge \cdots \wedge \theta^{i_k} \mid 1 \le i_1 < i_2 < \cdots < i_k \le n\}$$

is linearly independent and spans $\bigwedge^k(V^*)$, hence is a basis of $\bigwedge^k(V^*)$.

(b) Notation and hypotheses as in (a). Let W be another finite-dimensional vector space and let $A : V^* \to W^*$ be a linear map. In view of part (a), we can define a linear map "$\bigwedge^k(A; \mathcal{B})$" from $\bigwedge^k(V^*) \to \bigwedge^k(W^*)$ by setting

$$\bigwedge\nolimits^k(A; \mathcal{B})(\theta^{i_1} \wedge \cdots \wedge \theta^{i_k}) := A\theta^{i_1} \wedge \cdots \wedge A\theta^{i_k}$$

for each strictly increasing mutli-index $(i_1, \ldots, i_k)$, and extending linearly. Since the definition of $\bigwedge^k(A; \mathcal{B})$ relied explicitly on the choice of basis $\mathcal{B}$, it is reasonable to ask whether choosing a different basis $\mathcal{B}'$ leads to a different map $\bigwedge^k(A; \mathcal{B}')$.

To answer this question, show (**without using Proposition 4.13 or Corollary 4.14**) that

$$(\textstyle\bigwedge^k(A;\mathcal{B}))(\eta_1 \wedge \cdots \wedge \eta_k) = A\eta_1 \wedge \cdots \wedge A\eta_k \tag{4.38}$$

for all $(\eta_1, \ldots, \eta_k) \in (V^*)^k$. (Thus, in particular, this holds if the η_i are drawn from another basis $\mathcal{B}'$ of V^*.) ▲

5 Induced inner products

Throughout this section, V and W are vector spaces.

5.1 Some generalities about real-valued bilinear maps

In Section 2 we related the space $\mathrm{Bihom}(V \times W, \mathbf{R})$ (and its elements) to tensor products. Some general features of bilinear maps $V \times W \to \mathbf{R}$ for which tensor-products are not in the foreground, and were omitted from Section 2 for this reason, are discussed below.

Every bilinear map $B : V \times W \to \mathbf{R}$ canonically defines linear maps $\tilde{B}_1 : V \to W^*$ and $\tilde{B}_2 : W \to V^*$ by setting

$$\tilde{B}_1(v) = B(v, \cdot) : W \to \mathbf{R}, \tag{5.39}$$

$$\tilde{B}_2(w) = B(\cdot, w) : V \to \mathbf{R}. \tag{5.40}$$

For the rest of this subsection, given any bilinear map $B : V \times W \to \mathbf{R}$, the notation $\tilde{B}_1$, $\tilde{B}_2$ is as in (5.39)–(5.40).

Exercise 5.1 Show that if V and W are finite-dimensional, then each of the maps $\tilde{B}_1, \tilde{B}_2$ is the natural adjoint of the other, modulo the canonical identification of a finite-dimensional vector space with its double dual. (See Definition 6.6.)

Exercise 5.2 Let $T : \mathrm{Bihom}(V, W) \times V \times W \to \mathbf{R}$ be map defined by

$$T(B, v, w) = B(v, w).$$

Show that T is trilinear, and that the maps

$$\beta_1 : \mathrm{Bihom}(V \times W, \mathbf{R}) \to \mathrm{Hom}(V, W^*), \quad \beta_1(B) = \tilde{B}_1, \tag{5.41}$$

and

$$\beta_2 : \mathrm{Bihom}(V \times W, \mathbf{R}) \to \mathrm{Hom}(W, V^*), \quad \beta_2(B) = \tilde{B}_2, \tag{5.42}$$

are isomorphisms. ▲

There is some special terminology for bilinear maps $V \times W \to \mathbf{R}$ when V and W are the same vector space:

Definition 5.3 A *bilinear form* on a vector space V is a bilinear map $V \times V \to \mathbf{R}$. A *quadratic form* on V is a symmetric bilinear form.

Thus, an inner product is a positive-definite quadratic form.

...["kernel" for symmetric/antisymmetric bilinear forms on V]

Definition 5.4 For finite-dimensional vector spaces V and W, a bilinear map $B : V \times W \to \mathbf{R}$ is called *nondegenerate* if the maps $\tilde{B}_1 : V \to W^*$ and $\tilde{B}_2 : W \to V^*$ are isomorphisms.

A nondegenerate bilinear map $V \times W \to \mathbf{R}$ is also called a *perfect pairing* between the finite-dimensional vector spaces V and W.[7] ▲

Obviously, in the setting of Definition 5.4, a necessary condition for nondegeneracy of B is that $\dim(V) = \dim(W)$.

Two common examples of perfect pairings are the dual pairing $V^* \times V \to \mathbf{R}$ and an inner product $V \times V \to \mathbf{R}$ (for finite-dimensional V). For inner products, nondegeneracy follows from positive-definiteness; see Exercise 5.6 below.

Exercise 5.5 Assume that V and W are finite-dimensional, and let $B : V \times W \to \mathbf{R}$ be a blinear map. Show that $\tilde{B}_1$ is an isomorphism $\iff$ $\tilde{B}_2$ is an isomorphism.

Hence, in Definition 1, we could replace the condition that both maps $\tilde{B}_1, \tilde{B}_2$ are isomorphism by the condition that at least one of them is an isomorphism. ▲

Exercise 5.6 Let V and W be finite-dimensional vector spaces of equal dimension, and let $B : V \times W \to \mathbf{R}$ be a bilinear map. Show that B is nondegenerate iff for all $v \in V$,

$$B(v, w) = 0 \ \text{ for all } w \in W \quad \Longrightarrow \ v = 0; \tag{5.43}$$

equivalently, if $\ker(\tilde{B}_1) = \{0\}$.

Remark 5.7 The reason we have restricted attention to finite-dimensional vector spaces for most of this subsection is that in the infinite-dimensional case, even when B is an inner product on a vector space V, condition (5.43) is not enough to imply that the map $\tilde{B}_1$ ($= \tilde{B}_2$ since an inner product is symmetric) is an isomorphism. In the infinite-dimensional case a bilinear map $V \times W \to \mathbf{R}$ is called *weakly nondegenerate* if $\ker(\tilde{B}_1) = \{0\} = \ker(\tilde{B}_2)$.

[7] The "perfect pairing" terminology is seen more often when V and W are different vector spaces than when they are the same space.

5.2 [decide title]

Let $g : V \times V \to \mathbf{R},\ h : W \times W \to \mathbf{R}$ be bilinear maps. Then the map

$$\begin{aligned} V \times W \times V \times W &\to \mathbf{R}, \\ (v, w, v', w') &\mapsto g(v, v')h(w, w'), \end{aligned}$$

is multilinear, and hence determines a linear map $L : V \otimes W \otimes V \otimes W \to \mathbf{R}$. But $V \otimes W \otimes V \otimes W = (V \otimes W) \otimes (V \otimes W)$ (see (3.24) and Convention 3.11), so we may view L as a linear map $(V \otimes W) \otimes (V \otimes W) \to \mathbf{R}$. The map $\otimes_{\mathrm{op}} : (V \otimes W) \times (V \otimes W) \to (V \otimes W) \otimes (V \otimes W)$ (for this "$\otimes_{\mathrm{op}}$", replace each of the spaces V, W in Figure 1 with our currrent $V \otimes W$) then pulls L back to a bilinear map

$$g \otimes h = L \circ \otimes_{\mathrm{op}} : (V \otimes W) \times (V \otimes W) \to \mathbf{R}$$

(thus $g \otimes h \in \mathrm{Bihom}\big((V \otimes W) \times (V \otimes W), \mathbf{R}\big)$) satisfying

$$(g \otimes h)(v \otimes w, v' \otimes w') = g(v, v')\, h(w, w')$$

for all $v, v' \in V$ and $w, w' \in W$.

The notation "$g \otimes h$" warrants some explanation. By definition $g \in \mathrm{Bihom}(V \times V, \mathbf{R})$ and $h \in \mathrm{Bihom}(W \times W, \mathbf{R})$, so we may form the tensor product of g and h as an element of $\mathrm{Bihom}(V \times V, \mathbf{R}) \otimes \mathrm{Bihom}(W \times W, \mathbf{R})$. Let us temporarily denote this tensor product as $g \otimes' h$. There is a canonical linear injection $\mathrm{Bihom}(V \times V, \mathbf{R}) \otimes \mathrm{Bihom}(W \times W, \mathbf{R}) \hookrightarrow \mathrm{Bihom}\big((V \otimes W) \times (V \otimes W), \mathbf{R}\big)$ defined by the following composition:

$$\begin{aligned} \mathrm{Bihom}(V \times V, \mathbf{R}) \otimes \mathrm{Bihom}(W \times W, \mathbf{R}) \quad &\xrightarrow{\text{canon. iso.}} \quad (V \otimes V)^* \otimes (W \otimes W)^* \\ &\qquad\qquad \text{(see Proposition 2.25(a))} \\ &\xhookrightarrow{\text{canon. injection}} \quad \big((V \otimes V) \otimes (W \otimes W)\big)^* \\ &\qquad\qquad \text{(see Proposition 2.25(c))} \\ &\xrightarrow{\text{canon. iso.}} \quad \big((V \otimes W) \otimes (V \otimes W)\big)^* \\ &\qquad\qquad \text{(using associativity and Exercise 2.17)} \\ &\xrightarrow{\text{canon. iso.}} \quad \mathrm{Bihom}\big((V \otimes W) \times (V \otimes W), \mathbf{R}\big) \\ &\qquad\qquad \text{(see Proposition 2.25(a))}. \end{aligned}$$

This composition carries $g \otimes' h$ to the element we have denoted $g \otimes h \in \mathrm{Bihom}\big((V \otimes W) \times (V \otimes W), \mathbf{R}\big)$ (see Exercise 5.8 below). Since all the maps above are canonical, we are allowing ourselves to write the image of the "literal" tensor product $g \otimes' h$ simply as $g \otimes h$. Note also that if V and W are finite-dimensional, then by Proposition 2.30, all four maps in this composition are isomorphisms, so that $g \otimes h$ is the image of the literal tensor product of g and h under a canonical isomorphism.

Exercise 5.8 Check that in the argument above, the indicated composition carries $g \otimes' h$ to $g \otimes h$, as claimed.

Exercise 5.9 (a) Show that if each of the bilinear maps g, h is *symmetric*, then so is $g \otimes h$.

(b) Assume now that g and h are *inner products*: positive-definite as well as symmetric. You will show below that $g \otimes h$ is an inner product as well. By part (a), we already know that $g \otimes h$ is symmetric; the question is positive-definitness.

(i) Assume that $\{e_i\}_{i=1}^n \subset V$ and $\{f_i\}_{i=1}^m \subset W$ are, respectively a g-orthonormal and an h-orthonormal set (not necessarily bases; we are not assuming finite-dimensionality.) Show that $\{e_i \otimes f_j : 1 \leq i \leq n,\ 1 \leq j \leq m\}$ is a $(g \otimes h)$-orthonormal set.

(ii) Using part (i), Show that $g \otimes h$ is positive-definite, hence an inner product.

In view of the result of Exercise 5.9, it would be reasonable to call $g \otimes h$ a "tensor-product inner product". But because that terminology is so awkward, we will borrow differential geometers' habit of referring to an inner product as a *metric*:

Terminology (tensor-product metric). If g, h are inner products on V, W respectively, we refer to the inner product $g \otimes h$ on $V \otimes W$ as a *tensor-product metric*.

To emphasize: this terminology uses"metric" in the sense of "Riemannian metric at a point", not in the sense of "distance function" (i.e. not with the metric-space meaning).

6 Appendix: Linear Algebra "Review"

Some basic linear-algebraic facts that are often omitted from linear-algebra courses are "reviewed" here.

6.1 Bases

In any nontrivial vector space, whether finite- or infinite-dimensional, a *basis* $\mathcal{B}$ (in the purely algebraic sense) is defined to be a linearly independent spanning-set; equivalently, a maximal linearly independent set (where "maximal" means that for any $v \notin \mathcal{B}$, the set $\mathcal{B} \cup \{v\}$ is linearly dependent); equivalently, a minimal spanning set (where "minimal" means that if $v \in \mathcal{B}$, then $\mathcal{B} \setminus \{v\}$ does not span).[8]

[8]In a Hilbert space, and other topological vector spaces, sometimes "basis" is taken to mean a linearly independent $\mathcal{B}$ for which the *closure* of span($\mathcal{B}$) is the entire space, rather than requiring $\mathcal{B}$ itself to span the entire space. For finite-dimensional vector spaces, these two meanings of *basis* coincide, but for infinite-dimensional vector spaces. In the infinite-dimensional case, sometimes a basis in the purely algebraic sense is called a *Hamel basis*, to avoid any potential misinterpretation.

In order not to exclude the trivial vector space $\{0\}$ from various statements involving bases, it is conventional to define the empty set $\emptyset$ to be a basis of this space. Clearly $\emptyset$ is the only linearly independent set, hence is maximal among these. The "empty linear combination" is assigned the value 0, making $\emptyset$ a minimal spanning set for $\{0\}$.[9]

Remark 6.1 It is not obvious that an arbitrary infinite-dimensional vector space *has* a basis; for example, if you try to write down a basis of the space of all functions $\mathbf{R} \to \mathbf{R}$ (or something "tamer" like the space all continuous functions $[0,1] \to \mathbf{R}$, or the space $L^2([0,1])$), you are not likely to succeed, and you might convince yourself that a basis couldn't possibly exist. However, with the aid of Zorn's Lemma, one can quickly show that every nontrivial vector space does, indeed, have a basis; in fact, more generally any linearly independent subset of a vector space can be extended to a basis. (*Proof*: Let V be a vector space and let $\mathcal{B}$ be a linearly independent subset of a vector space V. Let $S = \{\mathcal{A} \subset S \mid \mathcal{A}$ is linearly independent and $\mathcal{B} \subset \mathcal{A}\}$. The set S is partially ordered by inclusion. Any chain $C \subset S$ has an upper bound, namely $\bigcup_{\mathcal{A} \in C} \mathcal{A}$. Hence, by Zorn's Lemma, S contains at least one maximal element.)

Remark 6.2 In these notes, we are using existence of bases of arbitrary vector spaces (with no assumption of finite-dimensionality) just to unify the presentation. A basis whose construction *requires* Zorn's Lemma (or anything else equivalent to the Axiom of Choice) is essentially useless for any *practical* purpose, or for providing an intuitive understanding of anything. For a finite-dimensional vector space, the usual, constructive proof of existence of a basis is far more useful and important than a proof that uses Zorn's Lemma. ▲

If $\mathcal{B}$ is a basis of the vector space V, then every $v \in V$ can be written uniquely as a linear combination of elements of $\mathcal{B}$; we may use the notation $v = \sum_{u \in B} a_u\, u$, with the understanding that this notation means the finite sum $\sum_{u \in \mathcal{B} \,|\, a_u \neq 0\}}$ and where $a_u \in \mathbf{R}$ for all $u \in V$.

Given vector spaces V and W, a basis $\mathcal{B}$ of V, and an arbitrary function $f : \mathcal{B} \to W$, we can extend f to a linear map $L : V \to W$, setting $L(\sum_{v \in \mathcal{B}} a_v\, v) := \sum_{v \in \mathcal{B}} a_v\, f(v)$. Since the representation of any given $x \in V$ as a linear combination of elements in $\mathcal{B}$ is unique, the linear map L defined this way is the *unique* extension of f to a *linear* map from V to W.

6.2 V^* and $\mathrm{Hom}(V, W)$

Let S be any nonempty set. As noted in Section 1, pointwise operations define a vector-space structure on the set $\mathrm{Func}(S, \mathbf{R})$. Similarly, for any vector space W, pointwise operations define a vector-space structure on the set $W^S := \{$all functions from S to $W\}$.

[9] A general convention for Σ-notation is that in any vector space V, any sum of the form $\sum_{\alpha \in A} v_\alpha$, where $v_\alpha \in V$ for all $\alpha \in A$, is assigned the value 0_V if $A = \emptyset$, and is called the *empty sum* in V.

Let V, W be vector spaces. In these notes, the notation $\text{Hom}(V, W) \subset W^V$ denotes the space of all linear maps from $V \to W$. As is easily checked, $\text{Hom}(V, W)$ is a subspace of W^V.

Definition 6.3 (Dual space) The *(algebraic) dual* of V is $V^* := \text{Hom}(V, \mathbf{R})$.

For $v \in V$ and $\theta \in V^*$, we define $\langle \theta, v \rangle := \theta(v)$. With this notation, the *dual pairing* $\langle \cdot, \cdot \rangle$ is a bilinear map $V^* \times V \to \mathbf{R}$. When more than one vector space is under discussion (e.g. V and W) we may use notation such as "$\langle \cdot, \cdot \rangle_V$" and "$\langle \cdot, \cdot \rangle_W$" for emphasis, even though context usually singles-out the only possible dual pairing we could be using in a given expression.

Exercise 6.4 Let V be a nontrivial vector space and let $v \in V$ be a nonzero element. Show that there exists $\theta \in V^*$ such that $\langle \theta, v \rangle = 1$.

Remark 6.5 Let V be a vector space. For each $v \in V$, define the *evaluation map* $\text{ev}_v : V^* \to \mathbf{R}$ by $\text{ev}_v(\theta) = \theta(v) = \langle \theta, v \rangle$. Clearly this evaluation map is linear, hence is an element of $V^{**} := (V^*)^*$. The map $\text{ev} : V \to V^{**}$ defined by $v \mapsto \text{ev}_v$ is easily seen to be linear and injective, hence can be thought of as a canonical inclusion of V into V^{**} (i.e. we can think of V as a subspace of V^{**} if we agree that, when we regard $v \in V$ as an element of V^{**}, we mean the element $\text{ev}_v \in V^{**}$).

In general, a linear map $A : V \to W$ does not determine a linear map $W \to V$. But *any* function $f : W \to X$ pulls back, via A, to a function $V \to X$, namely $f \circ A$. If X is a vector space and f is linear, then $f \circ A$ is linear as well. Hence A *does* determine a linear map $W^* = \text{Hom}(W, \mathbf{R}) \to V^* = \text{Hom}(V, \mathbf{R})$:

Definition 6.6 (natural adjoint of a linear map) Let $A \in \text{Hom}(V, W)$. The *natural adjoint* of A is the linear map $A^* : W^* \to V^*$ defined by

$$A^*\xi := A^*(\xi) = \xi \circ A. \tag{6.44}$$

Note that equation (6.44) can be written equivalently as

$$\langle A^*\xi, v \rangle_V = \langle \xi, Av \rangle_W \quad \text{for all } v \in V. \tag{6.45}$$

(The subscripts V and W could be omitted without causing any ambiguity, since on each side of the equation in (6.45), there is only one dual pairing that makes sense.)

[Note to self: I may want to change notation "A^*" later.]

Exercise 6.7 Let $A \in \text{Hom}(V, W)$ and let $A^* \in \text{Hom}(W^*, V^*)$ be the natural adjoint of A.

(a) Prove that if A is surjective, then A^* is injective.

(b) Prove that if A is injective, then A^* is surjective.

(c) Why does the combination of (a) and (b) not turn either statement into an "if and only if"?

Exercise 6.8 (a) Prove that the map $\mathrm{Hom}(V,W) \to \mathrm{Hom}(W^*,V^*)$ given by $A \mapsto A^*$ is linear.

(b) Let $A \in \mathrm{Hom}(V,W)$. Show that $A^{**} := (A^*)^* \in \mathrm{Hom}(V^{**},W^{**})$ "restricts" to A under the canonical inclusions $\mathrm{ev}_V : V \hookrightarrow V^{**}$, $\mathrm{ev}_W : W \hookrightarrow W^{**}$ (See Remark 6.5). I.e. show that $A^{**} \circ \mathrm{ev}_V = \mathrm{ev}_W \circ A$ (where $\mathrm{ev}_V(v)$ is what was called ev_v in Remark 6.5, etc. for W).

Exercise 6.9 Let V, W, Z be vector spaces and let $B : V \to W$ and $A : W \to Z$ be linear maps. Show that $(A \circ B)^* = B^* \circ A^*$.

Remark 6.10 Let $\mathcal{C}$ be the category whose objects are vector spaces and whose morphisms linear transformations. Exercise 6.9 shows that, in this setting, dualization is a contravariant functor $\mathcal{C} \to \mathcal{C}$. Under this functor, an object V is mapped to the dual space V^*, and a morphism A is mapped to the natural adjoint A^*.

Proposition 6.11 *For $f \in \mathrm{Func}(S,\mathbf{R})$, define $T_f : \mathbf{R}[S] \to \mathbf{R}$ by setting*

$$T_f\left(\sum_{p\in S} a_p\, p\right) := T_f\left(\sum_{p\in S} a_p\, e_p\right) = \sum_{p\in S} a_p f(p).$$

(In the first sum above we used "formal linear combination" notation. In the second sum we simply re-expressed the first sum using the notation (1.1)–(1.2), to prevent any misunderstanding of the notation in the first sum.)

1. *For all $f \in \mathrm{Func}(S,\mathbf{R})$, the map $T_f : \mathbf{R}[S] \to \mathbf{R}$ linear (hence an element of $(\mathbf{R}[S])^*$), and may be defined equivalently as the unique linear map $T' : \mathbf{R}[S] \to \mathbf{R}$ satisfying $T'(e_p) = f(p)$ for all $p \in S$.*

2. *The map $T : \mathrm{Func}(S,\mathbf{R}) \to (\mathbf{R}[S])^*$ defined by $T(f) = T_f$ is an isomorphism.*

Hence $(\mathbf{R}[S])^*$ is canonically isomorphic to $\mathrm{Func}(S,\mathbf{R})$.

Proof: Exercise. ■

Remark 6.12 In analysis, topological vector spaces (vector spaces equipped with a topology that renders the vector-space operations continuous) are of paramount importance. In that setting, the notation $\mathrm{Hom}(V,W)$ is often reserved for the space of *continuous* linear maps $V \to W$, and the notation V^* is used for the *continuous dual* of V, the space of *continuous* linear maps $V \to \mathbf{R}$.)

A *norm* on a vector space V determines a metric, and therefore a topology. For any *finite-dimensional* vector space, this topology is independent of the choice of norm (this is not true for infinite-dimensional vector spaces), and may thus be called "*the* norm topology" unambiguously. See Section 6.7.

The standard topology on a finite-dimensional vector space is the norm topology. When a finite-dimensional vector space is treated as a topological space without any topology having been specified explicity, the norm topology is being assumed implicitly.

All linear maps between *finite-dimensional* vector spaces are continuous (with respect to the norm topologies on the domain and codomain), so there is never any ambiguity in what the notation $\mathrm{Hom}(V, W)$ means in the finite-dimensional setting.

In general, infinite-dimensional vector spaces do *not* have a canonical topology, and the choice of topologies on domain and codomain affects which linear maps are continuous. If an infinite-dimensional vector space V is defined in such a way that V automatically acquires a particular norm, then generally the topology determined by that norm is assumed. But infinite-dimensional vector spaces admit inequivalent norms, defining different topologies. There are also infinite-dimensional topological vector spaces whose topology is not *any* norm topology. ▲

6.3 The dual of a basis

Let V be a nontrivial vector space and suppose that $\mathcal{B}$ is a basis of V. For each $v \in \mathcal{B}$, the function $e_v : \mathcal{B} \to \mathbf{R}$ defined in (1.1) extends to a unique linear map $\theta_v : V \to \mathbf{R}$. Thus, the linear map $\theta_v \in V^*$ satisfies $\theta_v(w) = \delta_{v,w}$ for all $v, w \in \mathcal{B}$. Let $\mathcal{B}' = \{\theta_v : v \in \mathcal{B}\}$. It is easily seen that $\mathcal{B}'$ is linearly independent. We may call $\mathcal{B}'$ the *subset of* V^* *dual to* $\mathcal{B}$.

If $\dim(V)$ is finite, it is easy to show that $\mathcal{B}'$ spans V^* (Exercise 6.13 below), hence is a basis of V^* (which also proves that $\dim(V^*) = \dim(V)$); we call $\mathcal{B}'$ the *basis of* V^* *dual to* $\mathcal{B}$ (or simply "the dual basis" if there is enough context to make it clear what basis of V this basis of V^* is dual to). If $\dim(V)$ is infinite, then in general $\mathcal{B}'$ will not span V^*.

Exercise 6.13 Let V be a finite-dimensional vector space and let $n = \dim(V)$.

(a) Check that if $V = \{0\}$ (equivalently, if $n = 0$), then $V^* = \{0\}$.

(b) Assume $n > 0$, let $\mathcal{B} := \{e_i\}_{i=1}^n$ be a basis of V, and let $\{\theta_i\}_{i=1}^n$ be the subset of V^* dual to $\mathcal{B}$. Show that the linearly independent set $\{\theta^i\}_{i=1}^n$ spans V^* by showing that each $\xi \in V^*$ satisfies $\xi = \sum_{i=1}^n a_i\,\theta^i$, where $a_i = \langle \xi, e_i \rangle$. Hence $\{\theta^i\}_{i=1}^n$ is a basis of V^*.

Deduce from (a) and (b) that $\dim(V^*) = \dim(V)$. ▲

Corollary 6.14 (of Exercise 6.13) *Suppose V is finite-dimensional. Then the canonical inclusion* $\mathrm{ev} : V \hookrightarrow V^{**}$ (see Remark 6.5) *is an isomorphism, and hence the spaces V and V^{**} are canonically isomorphic.*

Proof: By Exercise6.13, $\dim(V^{**}) = \dim(V^*) = \dim(V)$. Hence, by equidimensionality, the canonical injection $V \hookrightarrow V^{**}$ is an isomorphism. ■

Remark 6.15 Of course, by equidimensionality alone, a finite-dimensional vector space V is *isomorphic* to V^*, and V^* is isomorphic to V^{**}. The point of Corollary 6.14 is that the isomorphism from V to its *double*-dual is *canonical*, whereas, in general, there is no *canonical* isomorphism from V to its dual. (The "in general" matters. There is a large, important class of vector spaces for which there *is* a canonical isomorphism from V to its dual, namely finite-dimensional *inner-product* spaces. An inner product determines an isomorphism from a finite-dimensional vector space to its dual; see [later]. There are also other forms of additional structure, not just inner products, that can determine an isomorphism $V \to V^*$.)

Corollary 6.16 *Suppose V and W are finite-dimensional and let $A \in \mathrm{Hom}(V, W)$. Then $A^{**} = A$, modulo the canonical identifications of V^{**}, W^{**} with V, W respectively.*

Proof: Immediate from Exercise 6.8 and Corollary 6.14 and

Exercise 6.17 Let V be a vector space (not necessarily finite-dimensional) and suppose that $\{\xi^1, \ldots, \xi^n\}$ is a linearly independent set in V^*. *Without using Zorn's Lemma* (or anything else equivalent to the Axiom of Choice), show that there exist $v_1, \ldots, v_n \in V$ such that $\langle \xi^i, v_j \rangle = \delta^i_j$ for all $i, j \in \{1, \ldots, n\}$.

6.4 Finite-dimensional vector spaces: algebraic aspects

Let $M_{m\times n}(\mathbf{R})$ denote the space of $m \times n$ matrices. In keeping with the standard conventions for relating linear algebra to matrix algebra, we view elements of $\mathbf{R}^n$ as column vectors (equivalently, $n \times 1$ matrices)

With $\mathbf{R}^n$ identified with the space of n-component *column* vectors, matrix multiplication identifies the space of n-component *row* vectors (equivalently, $1 \times n$ matrices) with $(\mathbf{R}^n)^*$. The dual pairing $(\mathbf{R}^n)^* \times \mathbf{R}^n \to \mathbf{R}$ is given simply by $\langle\ th, v\rangle = \theta v$ (the product of a $1 \times n$ matrix θ on the left and an $n \times 1$ matrix v on the right). The matrix-transpose operation $M_{n\times 1}(\mathbf{R}) \to M_{1\times n}(\mathbf{R})$ is a (canonical) isomorphism $\mathbf{R}^n \to (\mathbf{R}^n)^*$, but this isomorphism relies heavily on extra structure that $\mathbf{R}^n$ has that a general n-dimensional vector space lacks. (In particular, $\mathbf{R}^n$ has a *canonical basis*, the standard basis.) The existence of this simple, explicit isomorphism $\mathbf{R}^n \to (\mathbf{R}^n)^*$, together with the fact that the standard inner product on $\mathbf{R}^n$ can be written as $(v, w) \mapsto v^T w$, is the source of many misunderstandings and inaccurate statements, some about inner products and some about tensors.

When we first learn linear algebra, we get into the habit of writing all indices as lower indices (subscripts): we write the components of $v \in \mathbf{R}^n$ as v_i; we write the entries of $A \in M_{m\times n}(\mathbf{R})$ in the form A_{ij} or a_{ij}, with the first index labeling the row, and the second index labeling the column. For $A \in M_{m\times n}(\mathbf{R})$ and $B \in M_{n\times k}(\mathbf{R})$, the product AB is given by $(AB)_{ij} = \sum_{l=1}^{n} A_{il}B_{lj}$. I.e. we sum over adjacent indices: the second index of A and the first index of B. The definition of Av for $v \in \mathbf{R}^n$ amounts to treating v as an $n \times 1$ matrix with entries $v_{_j} = v_{_{j1}}$. Similarly, the definition of (m-component row vector)$\times$ ($m \times n$ matrix) amounts to treating the row vector as a $1 \times m$ matrix whose entries all have 1 as the *first* index.

Some convenience, and even insight, can be gained if (i) we allow ourselves to express "scalar times vector" by writing the vector on the left and the scalar on the right, (ii) treat an (ordered) basis as a $1\times$(something) array of vectors, and (iii) extend the rule of "indices that we're going to sum over should preferably be adjacent" to expressions that involve such an array of vectors. For example, letting $\mathbf{e} := \{e_i\}_{i=1}^n$ and $\mathbf{e}' = \{e_i'\}_{i=1}^m$ denote the standard bases of $\mathbf{R}^n$ and $\mathbf{R}^m$ respectively, then for $A \in M_{m\times n}(\mathbf{R})$, the *components* of Av satisfy $(Av)_i = \sum_j A_{ij}\, v_j$, but the images of the *basis vectors* under the linear map $T_A : v \mapsto Av$ satisfy $T_A(e_j) = \sum_i A_{ij}e_i' = \sum_i e_i'\, A_{ij}$.[10]

Something crucial to the standard conventions relating *linear algebra* to *matrix algebra* is that, when dealing with matrices representing linear transformations, one should **never deviate from the rule that the first index labels the row and the second index labels the column.** However, it is actually *good* to deviate from the convention of putting both indices downstairs—as long as one deviates *correctly* (!), and maintains the distinction between what the first and second indices represent. (We discuss below how to do this deviation correctly, and why that deviation is *good*.) In each of the expressions A_{ij}, ${A^i}_j$ ${A_i}^j$, and A^{ij}, the first index is i and the second is j. The critical distinction between first and second indices is lost if one writes A_i^j or A_j^i.[11]

. . .

[transposes and natural adjoints]

. . .

[10]An insight that one can gain from this notation-convention: to be consistent with usual convention for the action of $\mathrm{GL}(n, \mathbf{R})$ on $\mathbf{R}^n$ (i.e. $(A, v) \mapsto Av$, a *left*-action), the preferred action of $\mathrm{GL}(n, \mathbf{R})$ on the set of *bases* of $\mathbf{R}^n$ is the *right*-action, $(A, \mathbf{e}) \mapsto \mathbf{e}A$. This manifests itself in the definition of "principal $GL(n, \mathbf{R})$-bundle"—or principal G-bundle, wiht G a subgroup of $GL(n, \mathbf{R})$. Conventionally we take the G-action on the bundle'to be a right-action.

[11]Fortunately, virtually no one uses notation like "A_j^i" for matrices of linear transformations. Unfortunately, for other tensors, especially those with more than two indices, some authors implicitly treat the index-order as being alphabetical. The problem with this should be self-evident if you ask yourself which index comes first in "A_3^7"—or "A_ξ^ζ" if you don't know the Greek alphabet *in order*. Although *in practice* this problem does not arise with matrices representing linear transformations, this problem that arises in more complicated or less familiar contexts can be distilled to this simple, familiar context.

There are *some* matrices for which it is immaterial which index comes first. For example, if we choose to write the Kronecker delta as "δ_j^i" rather than "δ_{ij}", we cannot get into trouble from the ambiguity in index-order.

Now let V, W be arbitrary vector spaces of finite, positive dimension n and m respectively, and let $\mathbf{e} := \{e_i\}_{i=1}^n$ and $\mathbf{f} := \{f_i\}_{i=1}^m$ be bases of V and W respectively.

. . . [adjoint w.r.t. an inner product]

6.5 Complements and direct sums

Recall that if V, W are vector spaces, the *direct sum*[12] $V \oplus W$ is defined to be the Cartesian product $V \times W$ together with following the zero-element and vector-space operations:

$$\begin{aligned} 0_{V\oplus W} &:= (0_V, 0_W). \\ (v,w)+(v',w') &:= (v+v', w+w') \text{ for all } (v,w),(v',w') \in V \times W \\ c(v,w) &:= (cv, cw) \text{ for all } c \in \mathbf{R} \text{ and } (v,w) \in V \times W. \end{aligned}$$

Remark 6.18 Many people use the notation "$V \times W$" for $V \oplus W$, as they may have been taught in a first course on linear algebra. This is not literally *wrong*; it is simply a convention that can lead to a great deal of confusion several settings, one of which is multilinear algebra (= tensor algebra).[13] For example, tensor product distributes over direct sum: given vector spaces V, W, and Z, there is a canonical isomorphism

$$V \otimes (W \oplus Z) \underset{\text{canon.}}{\cong} (V \otimes W) \oplus (V \otimes Z).$$

This looks much less natural if written as

$$V \otimes (W \times Z) \underset{\text{canon.}}{\cong} (V \otimes W) \times (V \otimes Z),$$

a situation only made worse by comparing it with another true statement,

$$V \otimes (W \otimes Z) \underset{\text{canon.}}{\cong} (V \otimes W) \otimes Z.$$

▲

Recall that, for subspaces X, Y of a vector space V, the notation $X + Y$ denotes $\{x + y : x \in X, y \in Y\}$, also written as $\text{span}\{X, Y\}$ and identical to $\text{span}\{X \cup Y\}$.

[12] Some people call this an "external" direct sum, to distinguish it from "internal" direct sum in which V and W start out as subspaces of a single larger space.

[13] Vector spaces are, among other things, (abelian) groups. For two general groups G and H, the notation $G \times H$ is used for the direct *product*. The direct product of G and H is the group whose underlying set is the Cartesian product $G \times H$, with group operations defined componentwise: $(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)$. When G and H are abelian, so is their direct product, and for each of $G, H, G \times H$, it is common to write the group operation additively rather than multiplicatively, leading to: $(g_1, h_1) + (g_2, h_2) = (g_1 + g_2, h_1 + h_2)$. For the same reason, when G and H are abelian, the direct *product* is often called the direct *sum*, and written $G \oplus H$.

Definition 6.19 *Let V be a vector space, $X \subset V$ a subspace. A subspace $Y \subset V$ is called a* complement *of X if (i) $X + Y = V$ and (ii) $X \cap Y = \{0\}$.* ▲

It is immediate that, in the context of Definition 6.19, Y is a complement of X iff X is a complement of Y. Hence we may call X and Y *(mutually) complementary*, or *complements of each other.*

Example 6.20 Let V, W be vector spaces. The subsets $V \times \{0_W\}$ and $\{0_V\} \times W$ are subspaces of $V \oplus W$. They are complements of each other and are canonically isomorphic to V and W respectively.

Remark 6.21 Note that the word "orthogonal" does not appear in Definition 6.19. Orthogonal complements are a *special case* of complements, and are not defined unless an inner product has been specified. Even then, in the absence of *completeness*, an "orthogonal complement" is not always a true *complement* in the sense of Definition 6.19.

In a *Hilbert* space (a complete inner-product space), the orthogonal complement $X^\perp$ of a *closed* subspace X always exists, is unique, and *is* a true complement. But a proper, nontrivial subspace of a mere *vector* space never has a *unique* complement; in fact it has *infinitely many* complements. For example, in $\mathbf{R}^2$, let X denote the subspace $\mathbf{R} \times \{0\}$, the "x-axis." Then *every* 1-dimensional subspace of $\mathbf{R}^2$ other than X itself—graphically, every non-horizontal straight line through the origin—is a complement of X. ▲

Remark 6.22 Every subspace of a vector space has a complement. To see this, let V be a vector space and let $X \subset V$ be a subspace. Let $\mathcal{B}_X$ be a basis of X. Extend $\mathcal{B}_X$ to a basis $\mathcal{B}$ of V (see Remark 6.1). Let $\mathcal{B}_Y = \mathcal{B} \setminus \mathcal{B}_X$ and let Let $Y = \text{span}(\mathcal{B}_Y)$. Then it is easily seen that Y is a complement of X.

However, in situations in which Zorn's Lemma (or anything equivalent) must be used to extend $\mathcal{B}_X$ to a basis of V, a complement Y constructed as above is "essentially useless"; see Remark 6.2. ▲

Recall that if V is a vector space and X is a subspace, then a (linear) *projection* of V onto X is a linear map $\pi : V \to X$ for which $\pi|_X = \text{id}_X$. We also call a linear map $P : V \to V$ a projection if $P^2 := P \circ P = P$; observe that any such P is a projection from V to image(P) in the previous sense of "projection". Note that given a projection $\pi : V \to X$ as in the first sense of "projection", if $\iota : X \hookrightarrow V$ denotes the natural inclusion map of X into V, then the map $P = \iota \circ \pi : V \to V$ is a projection in the the second sense, a linear map $P \to P$ satisfying $P^2 = P$.

When X and Y are complements of each other in V, we also say that V is the *direct sum* of these two subspaces. Parts (c) and (d) of the next proposition express how the two uses of the term "direct sum" are related.

Proposition 6.23 *Let V be a vector space and let $X, Y \subset V$ be mutually complementary subspaces. Then:*

(a) For each $v \in V$, there exist unique *elements $v_X \in X, v_Y \in Y$ such that $v = v_X + v_Y$.*

(b) Because of the uniqueness in (a), there are well defined maps $\pi_X : V \to X$, $\pi_Y : V \to Y$, satisfying $v = \pi_X(v) + \pi_Y(v)$ for all $v \in V$. These maps are (linear) projections. If $\iota_X : X \hookrightarrow V$ and $\iota_Y : Y \hookrightarrow V$ the inclusion maps of X and Y, respectively, into V, and we define $\tilde{\pi}_X = \iota_X \circ \pi_X$ and $\tilde{\pi}_Y = \iota_Y \circ \pi_Y$, then $\tilde{\pi}_X$ and $\tilde{\pi}_Y$ are projection maps $V \to V$ satisfying $\tilde{\pi}_X + \tilde{\pi}_Y = \mathrm{id}_V$.

(c) The map $L : X \oplus Y \to V$ defined by $L(x, y) = x + y$ is an isomorphism.

(d) Define $\mathrm{proj}_1 : X \oplus Y \to X$ *and* $\mathrm{proj}_2 : X \oplus Y \to Y$ *by* $\mathrm{proj}_1(x, y) = x$ *and* $\mathrm{proj}_2(x, y) = y$. *Then the following diagrams commute.*

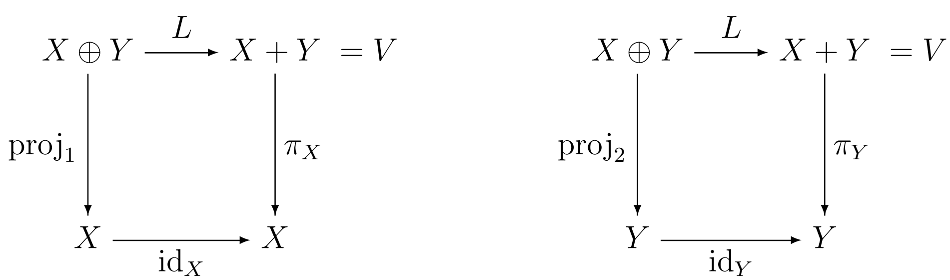

Figure 6: Diagrams for Proposition 6.23

Remark 6.24 Note that a subspace X of a vector space V does not determine a projection-map $V \to X$ all by itself. In Proposition 6.23, changing the choice of complement Y changes the map π_X. *Both* complementary subspaces X and Y are needed in order to define *either* projection-map π_X, π_Y.

In a Hilbert space, the notion of *orthogonal projection onto a closed subspace X* makes implicit use of the uniqueness of the *orthogonal* complement $X^\perp$; see Remark 6.21.▲

Remark 6.25 In the setting of *topological* vector spaces, the decomposition of a vector space V as the direct sum of two subspaces X and Y is useful only if the projection maps π_X, π_Y in Proposition 6.23 are *continuous.* A necessary condition for the continuity of these projections is that the subspaces X and Y be *closed.* (As an extreme example, suppose that X is a dense, proper subspace of V, and let Y be a complement of X. By taking a sequence in X converging to a nonzero element of Y, it is easy to see that the projection π_Y cannot be continuous, and therefore neither can $\pi_X = \mathrm{id}_V - \pi_Y$.)

6.6 Quotient vector spaces

Throughout this subsection, V is a vector space and $H \subset V$ a subspace.

A vector space is a module over $\mathbf{R}$, and a subspace is a submodule. Hence the *quotient (vector) space* V/H is defined; it is simply the quotient $\mathbf{R}$-module. Thus the elements of V/H are the H-cosets in V (translates of H by elements of V). The zero element of V/H is the coset $H = 0_V + H$, and the vector-space operations on V/H are defined by:

$$\begin{aligned} (v_1 + H) + (v_2 + H) &:= (v_1 + v_2) + H \quad \text{for all } v_1, v_2 \in V, \\ c(v + H) &:= (cv) + H \quad \text{for all } c \in \mathbf{R},\ v \in V. \end{aligned}$$

(Students not familiar with the definition of quotient modules should check that the operations above are well-defined. Well-definedness is something that needs to be checked—unless one has done it before—since, except in the case $H = \{0\}$, a given element $x \in V/H$ does not *uniquely* determine an element $v \in V$ for which $x = v+H$.)

Exercise 6.26 Show that, if V is finite-dimensional, then

$$\dim(V/H) = \dim(V) - \dim(H).$$

Remark 6.27 As a *set*, a quotient vector space V/H is simply $V/\sim$, where $\sim$ is the equivalence relation on V defined by declaring $v_1 \sim v_2 \iff v_1 - v_2 \in H$.

Proposition 6.28 *Let V be a vector space, let H be a subspace of V, and let Y be a complement of H. Let $\pi_H : V \to H$ and $\pi_Y : V \to Y$ be the projection maps defined in Proposition 6.23(b). Let $\pi : V \to V/H$ denote the quotient map. Then $\pi|_Y : Y \to V/H$ is an isomorphism.*

Thus, <u>for any given complement</u> Y of H, the quotient *space V/H is canonically isomorphic to the* sub*space Y.*

Proof: Exercise. ■

Remark 6.29 Do not be misled by Proposition 6.28 into thinking that a quotient space is "the same thing as" complementary subspace. The quotient space V/H is *canonically* defined by the pair (V, H). As noted earlier, in general a subspace H has *infinitely many* complements. Additional structure (e.g. a Hilbert space structure) is needed if one wishes to single out a *preferred* complement.

6.7 Finite-dimensional vector spaces: topological aspects

As mentioned in Remark 6.12, all norms on a given finite-dimensional vector space determine the same topology (called the *norm topology*). This topological independence-of-norm property is a consequence of the following fundamental result (usually proven in an advanced calculus course).

Theorem 6.30 *On any finite-dimensional vector space V, all norms are equivalent. I.e. given any two norms $\| \ \|$, $\| \ \|'$ on V, there exist constants c_1, c_2 such that $\|v\|' \leq c_1\|v\|$ and $\|v\| \leq c_2\|v\|'$.*

The norm topology is the standard topology on any finite-dimensional vector space, and is the "default" topology assumed for the remainder of this subsection.

Exercise 6.31 Show that every subspace of a finite-dimensional vector space is closed.

Theorem 6.30 has many important corollaries, some of which are assembled into a single multi-part corollary below.

Corollary 6.32 *Let V, W be finite-dimensional vector spaces.*

1. *Every linear map $V \to W$ is continuous.*
2. *The product topology on $V \times W$ coincides with the norm topology on $V \oplus W$.*
3. *If $W \subset V$, then the relative topology on W (as a subset of V) coincides with the norm topology on W.*
4. *Every surjective linear map $V \to W$ is an open map.*
5. *If $W \subset V$, the quotient topology on V/W coincides with the norm topology. Hence, in the finite-dimensional setting, the isomorphism in Proposition 6.28 is a homeomorphism with respect to the norm topology on the domain and the quotient topology on the codomain.*

Exercise 6.33 Prove Corollary 6.32