Chapter V. Tensors and Multilinear Forms

V.1. Multilinear Forms

Note. We generalize the idea of bilinear forms to the idea of multilinear forms. We define the tensor product of vector spaces and tensors, which are elements of a tensor product of copies of a vector space and copies of its dual space.

Definition V.1.01. A function $\mathbf{f} : X_1 \times X_2 \times \cdots \times X_n \to Y$ where $X_1, X_2, \ldots, X_n, Y$ are vector spaces is a *multilinear* mapping if

(i) $\mathbf{f}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i + \mathbf{x}_i', \ldots, \mathbf{x}_n) = \mathbf{f}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n) + \mathbf{f}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i', \ldots, \mathbf{x}_n)$

(ii) $\mathbf{f}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}a, \ldots, \mathbf{x}_n) = \mathbf{f}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)a$

for all $\mathbf{x}_1 \in X_1$, $\mathbf{x}_2 \in X_2$, $\ldots$, $\mathbf{x}_n \in X_n$, $\mathbf{x}_i' \in X_i$ for $i \in \{1, 2, \ldots, n\}$, and $a \in \mathbb{R}$. The collection of all such functions (which is shown to be a vector space is Exercise V.1.1) is denoted $L(X_1, X_2, \ldots, X_n; Y)$. We denote $L(\underbrace{X, X, \ldots, X}_{n \text{ times}}; Y)$ as $L^n(X; Y)$. If $\mathbf{f} \in L^n(X; \mathbb{R})$ then $\mathbf{f}$ is called a *multilinear form* on X.

Example V.1.02(iii). Let X be a vector space. Define $\mathbf{f} : X \times X^* \to \mathbb{R}$ as $\mathbf{f}(\mathbf{x}, \mathbf{g}) = \mathbf{g}(\mathbf{x})$. Then $\mathbf{f}$ is linear in the first variable by the linearity of each $\mathbf{g} \in X^*$ and is linear in the second variable by the definition of addition and scalar multiplication in X^*. Then $\mathbf{f} \in L(X, X^*; \mathbb{R})$.

Example V.1.02(iv). Let X be a vector space and let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in X$. Define $\mathbf{f} : X^* \times X^* \times \cdots \times X^* \to \mathbb{R}$ as $\mathbf{f}(\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n) = \mathbf{g}_1(\mathbf{x}_1)\mathbf{g}_2(\mathbf{x}_2)\cdots\mathbf{g}_n(\mathbf{x}_n)$. Then $\mathbf{f} \in L^n(X^*; \mathbb{R})$.

Note. Let X be a vector space and $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n \in X^*$. Define $\mathbf{g} : X \times X \times \cdots \times X \to \mathbb{R}^n$ as $\mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \mathbf{g}_1(\mathbf{x}_1)\mathbf{g}_2(\mathbf{x}_2)\cdots\mathbf{g}_n(\mathbf{x}_n)$. Then $\mathbf{g} \in L^n(X; \mathbb{R})$. We have $\mathbf{G}$ multilinear and want to think of $\mathbf{g}$ as a product of $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n$. This is where we encounter the tensor product.

Definition. Let X be a vector space and let $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n \in X^*$. The *tensor product* of $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n$ is $\mathbf{g} : X \times X \times \cdots \times X \to \mathbb{R}$ defines as $\mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \mathbf{g}_1(\mathbf{x}_1)\mathbf{g}_2(\mathbf{x}_2)\cdots\mathbf{g}_n(\mathbf{x}_n)$. We denote the tensor product as $\mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdot \otimes \mathbf{g}_n$.

Note/Definition. Since each $\mathbf{g}_i$ is linear, then $\mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdots \otimes \mathbf{g}_n$ is multilinear and so is in $L^n(X; \mathbb{R})$. Similarly, if $\mathbf{g}_1 \in X_1^*$, $\mathbf{g}_2 \in X_2^*$, $\ldots$, $\mathbf{g}_n \in X_n^*$ then we can define the *tensor product* $\mathbf{g} = \mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdot \otimes \mathbf{g}_n$ in $L(X_1, X_2, \ldots, X_n; \mathbb{R})$ as $\mathbf{g}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \mathbf{g}_1(\mathbf{x}_1)\mathbf{g}_2(\mathbf{x}_2)\cdots\mathbf{g}_n(\mathbf{x})$. So there is a mapping, which we denote as $\bigotimes$ such that $\bigotimes : X_1^* \times X_2^* \times \cdots \times X_n^* \to L(X_1, X_2, \ldots, X_n; \mathbb{R})$ defined as $\bigotimes((\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n)) = \mathbf{g} = \mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdots \otimes \mathbf{g}_n$. In Exercise V.1.3 it is to be shown that $\bigotimes$ is multilinear.

Note. We now consider a specific example where $X_1 = X_2 = \mathbb{R}^2$ so that $\bigotimes : (\mathbb{R}^2)^* \times (\mathbb{R}^2)^* \to L(\mathbb{R}^2, \mathbb{R}^2; \mathbb{R}) = L^2(\mathbb{R}^2; \mathbb{R})$ and discuss some shortcomings of this mapping.

Note. Let $\{\mathbf{b}_1, \mathbf{b}_2\}$ be any basis of $\mathbb{R}^2$. Then for $\mathbf{f} \in L^2(\mathbb{R}^2; \mathbb{R})$ a bilinear form on $\mathbb{R}^2 \times \mathbb{R}^2$ we have for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$:

$$\begin{aligned}
\mathbf{F}(\mathbf{x}, \mathbf{y}) &= \mathbf{F}((x^1, x^2), (y^1, y^2)) = \mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2, y^1\mathbf{b}_1 + y^2\mathbf{b}_2) \\
&= x^1y^2\mathbf{F}(\mathbf{b}_1, \mathbf{b}_2) + x^1y^2\mathbf{F}(\mathbf{b}_1, \mathbf{b}_2) + x^2y^1\mathbf{F}(\mathbf{b}_2, \mathbf{b}_1) + x^2y^2\mathbf{F}(\mathbf{b}_2, \mathbf{b}_2) \\
&\quad \text{by the bilinearity of } \mathbf{F} \\
&= x^iy^j\mathbf{F}(\mathbf{b}_i, \mathbf{b}_j) \text{ with the Einstein summation convention.}
\end{aligned}$$

Setting $\mathbf{F}(\mathbf{b}_i, \mathbf{b}_j) = f_{ij}$ we have

$$\begin{aligned}
\mathbf{F}(\mathbf{x}, \mathbf{y}) &= f_{ij}x^iy^j \\
&= f_{ij}(\mathbf{b}^i(\mathbf{x}))(\mathbf{b}^j(\mathbf{y})) \text{ by the definition of} \\
&\quad \text{the dual basis } \{\mathbf{b}^1, \mathbf{b}^2\} \text{ of } X^* = (\mathbb{R}^2)^* \\
&= f_{ij}(\mathbf{b}^i \otimes \mathbf{b}^j(\mathbf{x}, \mathbf{y})),
\end{aligned}$$

and hence $\mathbf{F} = f_{ij}\mathbf{b}^i \otimes \mathbf{b}^j$. Since $\mathbf{F}$ is an arbitrary element of $L^2(\mathbb{R}^2; \mathbb{R})$ and $\mathbf{F}$ is a linear combination of $\{\mathbf{b}^1 \otimes \mathbf{b}^1, \mathbf{b}^1 \otimes \mathbf{b}^2, \mathbf{b}^2 \otimes \mathbf{b}^1, \mathbf{b}^2 \otimes \mathbf{b}^2\}$ then this is a spanning set of $L^2(\mathbb{R}^2; \mathbb{R})$.

Note. Notice that $\mathbf{b}^1 \otimes \mathbf{b}^2 \neq \mathbf{b}^2 \otimes \mathbf{b}^1$ since, for example, with $(\mathbf{x}, \mathbf{y}) = (\mathbf{b}_1 + 2\mathbf{b}_2, 3\mathbf{b}_1 + 4\mathbf{b}_2) \in \mathbb{R}^2 \otimes \mathbb{R}^2$ we have

$$\mathbf{b}^1 \otimes \mathbf{b}^2 = \mathbf{b}^1_1(\mathbf{x})\mathbf{b}^2(\mathbf{y}) = x^1y^2 = (1)(4) = 4$$

$$\neq \mathbf{b}^2 \otimes \mathbf{b}^1(\mathbf{x}, \mathbf{y}) = \mathbf{b}^2(\mathbf{x})\mathbf{b}^1(\mathbf{y}) = x^2 1^1 = (2)(3) = 6.$$

Hence tensor product are not in general commutative.

Note/Definition. Though a sum of tensor products in $(\mathbb{R}^2)^* \times (\mathbb{R}^2)^*$ is an element of $L^2(\mathbb{R}^2; \mathbb{R})$, the sum may not itself be a tensor product. More generally, for $\mathbf{g}_1 \in X_1^*, \mathbf{g}_2 \in X_2^*, \ldots, \mathbf{g}_n \in X_n^*$ then we call $\mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdots \otimes \mathbf{g}_n \in L(X_1, X_2, \ldots, X_n; \mathbb{R})$ a *simple tensor* (or a *pure tensor*).

Note. If $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis for X_1 and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for X_2, then with $\{\mathbf{u}^1, \mathbf{u}^2\}$ and $\{\mathbf{v}^1, \mathbf{v}^2\}$ are the dual bases of X_1^* and X_2^* then we claim that if

$$a_{11}\mathbf{u}^1 \otimes \mathbf{v}^1 + a_{12}\mathbf{u}^1 \otimes \mathbf{v}^2 + 2_{21}\mathbf{u}^2 \otimes \mathbf{v}^1 + a_{22}\mathbf{u}^2 \otimes \mathbf{v}^2$$

is a simple tensor the $a_{11}a_{22} - a_{12}a_{21} = 0$. For a simple tensor $\mathbf{g}$, we have $\mathbf{g} = (a\mathbf{u}^1 + b\mathbf{u}^2) \otimes (c\mathbf{v}^1 + d\mathbf{v}^2)$ and notice that for $\mathbf{x} = x^1\mathbf{u}_1 + x^2\mathbf{u}_2 \in X_1$ and $\mathbf{y} = y^1\mathbf{v}^1 + y^2\mathbf{v}_2 \in X_2$ we have

$$\mathbf{g}(\mathbf{x}, \mathbf{y}) = ((a\mathbf{u}^1 + b\mathbf{u}^2)(\mathbf{x}))((c\mathbf{v}^1 + b\mathbf{v}^2)(\mathbf{y})) = (a\mathbf{u}^1(\mathbf{x}) + b\mathbf{u}^2(\mathbf{x}))(c\mathbf{v}^1(\mathbf{y}) + d\mathbf{v}^2(\mathbf{y}))$$

$$= ac\mathbf{u}^1 \otimes \mathbf{v}^1(\mathbf{x}, \mathbf{y}) + ad\mathbf{u}^1 \otimes \mathbf{v}^2(\mathbf{x}, \mathbf{y}) + bc\mathbf{u}^2 \otimes \mathbf{v}^1(\mathbf{x}, \mathbf{y}) + bd\mathbf{u}^2 \otimes \mathbf{v}^2(\mathbf{x}, \mathbf{y})$$

$$= (ac\mathbf{u}^1 \otimes \mathbf{v}^1 + ad\mathbf{u}^1 \otimes \mathbf{v}^2 + bc\mathbf{u}^2 \otimes \mathbf{v}^1 + bd\mathbf{u}^2 \otimes \mathbf{v}^2)(\mathbf{x}, \mathbf{y}),$$

so that

$$(a\mathbf{u}^1 + b\mathbf{u}^2) \otimes (c\mathbf{v}^1 + d\mathbf{v}^2) = ac\mathbf{u}^1 \otimes \mathbf{v}^1 + ad\mathbf{u}^1 \otimes \mathbf{v}^2 + bc\mathbf{u}^2 \otimes \mathbf{v}^1 + bd\mathbf{u}^2 \otimes \mathbf{v}^2$$

$$= a_{11}\mathbf{u}^1 \otimes \mathbf{v}^1 + a_{12}\mathbf{u}^1 \otimes \mathbf{v}^2 + a_{21}\mathbf{u}^2 \otimes \mathbf{v}^1 + a_{22}\mathbf{u}^2 \otimes \mathbf{v}^2 \qquad (*)$$

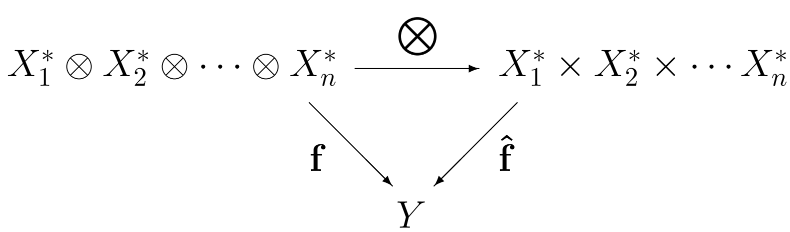
where $a_{11} = ac$, $a_{12} = ad$, $a_{21} = bc$, and $a_{22} = bd$. Notice that $a_{11}a_{22} - a_{12}a_{21} = 0$. So if $(*)$ is a simple tensor then $a_{11}a_{22} - a_{12}a_{21} = 0$. Hence, if $a_{11}a_{22} - a_{12}a_{21} \neq 0$ then $(*)$ does not represent a simple tensor. For example, with $a_{11} = a_{22} = 1$ and $a_{12} = a_{21} = 0$ we see that $\mathbf{u}^1 \otimes \mathbf{v}^1 + \mathbf{u}^2 \otimes \mathbf{v}^2$ is not a simple tensor. So the collection of simple tensors is not closed under addition and so is not a vector space. In particular, the collection of simple tensors in $L^2(\mathbb{R}^2;\mathbb{R})$ is not a vector space.

Note. We now see that we cannot take the tensor product of spaces X_1^*, X_2^*, $\ldots$, X_n^* simply as the collection of tensor products $\mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdots \otimes \mathbf{g}_n$ since these do not form a vector space. But we have also seen that, since $\{\mathbf{b}^1 \otimes \mathbf{b}^1, \mathbf{b}^1 \otimes \mathbf{b}^2, \mathbf{b}^2 \otimes \mathbf{b}^1, \mathbf{b}^2 \otimes \mathbf{b}^2\}$ is a spanning set for $L^2(\mathbb{R}^2;\mathbb{R})$, that any linear combination of simple tensors is again in $L^2(\mathbb{R}^2;\mathbb{R})$. So a "good candidate" (as Dodson and Poston put it on page 101) for the tensor product of space X_1^*, X_2^*, $\ldots$, X_n^* is $L(X_1, X_2, \ldots, X_n;\mathbb{R})$. In Exercise V.1.4(b) and V.1.4(d) we will show that the following hold.

(T i) $\bigotimes : X_1^* \times X_2^* \times \cdots \times X_n^* \to X_1^* \otimes X_2^* \otimes \cdots \otimes X_n^*$ is multilinear.

(T ii) If $\mathbf{f} : X_1^* \times X_2^* \times \cdots \times X_n^* \to Y$ (where Y is some vector space) is multilinear, then there is a unique linear map $\hat{\mathbf{f}} : X_1^* \otimes X_2^* \otimes \cdots \otimes X_n^* \to Y$ such that $\mathbf{f} = \hat{\mathbf{f}} \circ \bigotimes$.

Property (T ii) implies that the following diagram commutes:



In fact, these two properties are sufficient to define a tensor product of spaces up to isomorphism. We now shift notation and denote the spaces as $X_1, X_2, \ldots, X_n$.

Definition V.1.04. A *tensor product* of vector spaces $X_1, X_2, \ldots, X_n$ is a vector space X together with a map $\bigotimes : X_1 \times X_2 \times \cdots \times X_n \to X$ which satisfies properties (T i) and (T ii) above.

Lemma V.1.05. A tensor product of finite dimensional vector spaces $X_1, X_2, \ldots, X_n$ always exists and any two are isomorphic "in a natural way."

Note. In light of Lemma V.1.05, we speak of *the* tensor product of vector spaces $X_1, X_2, \ldots, X_n$ and denote this as $X_1 \otimes X_2 \otimes \cdots \otimes X_n$.

Note. The elements of $X_1 \otimes X_2 \otimes \cdots \otimes X_n$ are linear combinations of simple tensors:

$$X_1 \otimes X_2 \otimes \cdots \otimes X_n = \{(\mathbf{x}_{11} \otimes \mathbf{x}_{12} \otimes \cdots \otimes \mathbf{x}_{1n} \mid \mathbf{x}_{im} \in X_1$$

$$\text{and } a_i \in \mathbb{R} \text{ for } 1 \leq i \leq k \text{ and } 1 \leq m \leq n, k \in \mathbb{N}\}.$$

The following is to be proved in Exercise V.1.5.

Lemma V.1.A. Let $X_1, X_2, \ldots, X_n$ be vector spaces. Let $\mathbf{x}_i, \mathbf{x}'_i \in X_i$ for $1 \leq i \leq n$ and let $a \in \mathbb{R}$. Then

(T A) $\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \cdots \otimes (\mathbf{x}_i + \mathbf{x}'_i) \otimes \cdots \otimes \mathbf{x}_n = \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \cdots \otimes \mathbf{x}_i \otimes \cdots \otimes \mathbf{x}_n = \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \cdots \otimes \mathbf{x}'_i \otimes \cdots \otimes \mathbf{x}_n$.

(T S) $(\mathbf{x}_1 a) \otimes \mathbf{x}_2 \otimes \cdots \otimes \mathbf{x}_n = \mathbf{x}_1 \otimes (\mathbf{x}_2 a) \otimes \cdots \otimes \mathbf{x}_n = \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \cdots \otimes (\mathbf{x}_n a) = (\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \cdots \otimes \mathbf{x}_n)a$.

Note. Dodson and Poston state that we could use (T A) and (T S) to characterize the elements of $X_1 \otimes X_2 \otimes \cdots \otimes X_n$, "but this would involve more definitions." They then defend the fact that (T i) and (T ii) characterize the tensor product of vector spaces and (T A) and (T S) characterize the properties of the tensor product of vectors. They defend this view with the isomorphism conclusion of Lemma V.1.05 (see page 103).

Lemma. V.1.07. There is a "natural" isomorphism yielding $X_1^* \otimes X_2^* \otimes \cdots \otimes X_n^* \cong (X_1 \otimes X_2 \otimes \cdots \otimes X_n)^*$.

Note. Dodson and Poston claim that Lemma V.1.07 illustrates the usefulness of the tensor product. It "reduces the theory of multilinear forms on a collection of spaces [i.e., elements of $L(X_1, X_2, \ldots, X_n; \mathbb{R}) = X_1^* \otimes X_2^* \otimes \cdots \otimes X_n^*)$] to that of linear functionals on a single space [i.e., elements of $X_1 \otimes X_2 \otimes \cdots \otimes X_n$]." See page 104.

Lemma V.1.08. For any two vector spaces X_1 and X_2, there is a "natural" isomorphism yielding $L(X_1; X_2) \cong X_1^* \otimes X_2$.

Note. At this stage, we have defined

1. The tensor product of covariant vectors from the same space $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n \in X^*$,

$\mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdots \otimes \mathbf{g}_n$,

2. the tensor product of covariant vectors from different spaces $\mathbf{g}_1 \in X_1^*$, $\mathbf{g}_2 \in X_2^*$, $\ldots$, $\mathbf{g}_n \in X_n^*$, $\mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \cdots \otimes \mathbf{g}_n$.

3. the tensor product of (covariant) vector spaces $X_1^*, X_2^*, \ldots, X_n^*$ as $X_1^* \otimes X_2^* \otimes \cdots \otimes X_n^* = L(X_1, X_2, \ldots, X_n; \mathbb{R})$, and

4. the tensor product of vector spaces $X_1, X_2, \ldots, X_n$ (in Definition V.1.04).

We now define the tensor product of linear maps $\mathbf{A}_i : X_i \to Y_i$ for $1 \le i \le n$ where X_i and Y_i are vector spaces.

Note. Let $\mathbf{A}_i : X_i \to Y_i$ be linear maps for $1 \le i \le n$. Then we have the mappings

$$X_1 \times X_2 \times \cdots \times X_n \stackrel{(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)}{\longrightarrow} Y_1 \times Y_2 \times \cdots \times Y_n \stackrel{\otimes}{\longrightarrow} Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$$

where $(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = (\mathbf{A}_1\mathbf{x}_2, \mathbf{A}_2\mathbf{x}_2, \ldots, \mathbf{A}_n\mathbf{x}_n)$ and

$$\bigotimes((\mathbf{A}_1\mathbf{x}_1, \mathbf{A}_2\mathbf{x}_2, \ldots, \mathbf{A}_n\mathbf{x}_n)) = (\mathbf{A}_1\mathbf{x}_1) \otimes (\mathbf{A}_2\mathbf{x}_2) \otimes \cdots \otimes (\mathbf{A}_n\mathbf{x}_n).$$

Let $\mathbf{h} : X_1 \times X_2 \times \cdots \times X_n \to Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$ denote the composition $\bigotimes \circ (\mathbf{A}_1 \mathbf{A}_2, \ldots, \mathbf{A}_n)$.

Lemma V.1.B. Let $\mathbf{A}_i : X_i \to Y_i$ be linear maps for $1 \le i \le n$. Mapping $\mathbf{h} : X_1 \times X_2 \times \cdots \times X_n \to Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$ defined as $\mathbf{h} = \bigotimes \circ (\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)$ is multilinear.

Note. Lemma V.1.B shows that (T i) holds (with X_i^* of (T i) replaced with X_i here). By (T ii) (and its proof in Exercise V.1.4(d)) with $Y = Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$ and X_i^* of (T ii) replaced with X_i, and $\mathbf{h} : X_1 \times X_2 \times \cdots \times X_n \to Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$, there is a unique linear $\hat{\mathbf{h}} : X_1 \otimes X_2 \otimes \cdots \otimes X_n \to Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$ such that $\mathbf{h} = \hat{\mathbf{h}} \circ \bigotimes$ (where X_j^* of (T ii) is replaced with X_i here, and Y of (T ii) is replaced with $Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$ here).

Definition. Let X_i and Y_i be vector spaces and let $\mathbf{A}_i : X_i \to Y_i$ be linear maps for $1 \le i \le n$. The unique linear map $\hat{\mathbf{h}} : X_1 \times X_2 \times \cdots \times X_n \to Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$, such that $\mathbf{h} = \hat{\mathbf{h}} \circ \bigotimes$ where $\mathbf{h} : X_1 \times X_2 \times \cdots \times X_n \to Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n$ defined as $\mathbf{h} = \bigotimes \circ (\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)$, is the *tensor product* of maps $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$, denoted $\hat{\mathbf{h}} = \mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \cdots \otimes \mathbf{A}_n$.

Note/Definition. We now shift gears and consider tensor products of copies of a vector space X and its dual X^*. In particular, we are interested in the tensor product of spaces

$$\underbrace{X \otimes X \otimes \cdots \otimes X}_{k \text{ times}} \otimes \underbrace{X^* \otimes X^* \otimes \cdots \otimes X^*}_{h \text{ times}}.$$

This is denoted X_h^k. Vectors in X_h^k are called *tensors of X, covariant of degree h and contravariant of degree k*, or of *type* $\binom{k}{h}$. We denote $X_0^k = X^k$ and $X_h^0 = X_h$ (so that $X = X^1$ and $X^* = X_1$). By convention, we take $X_0^0 = \mathbb{R}$. By Exercise V.1.4(c), $\dim(X_h^k) = (\dim(X))^{k+h}$.

Note. We could have a tensor product of copies of X and X^* but not have all of the X's first followed by all of the X^*'s. Though we can permute the order in a tensor product of vector spaces to produce an isomorphic tensor product (by Exercise V.1.8(b)), we will often leave the product in the original order. For example,

$$X \otimes X \otimes X \otimes X^* \otimes X^* \otimes X \otimes X \otimes X^* \otimes X \otimes X$$

is denoted $X^{3}{}_{2}{}^{2}{}_{1}{}^{2}$ and its elements are covariant of degree $2+1=3$, contravariant of degree $3+2+2=7$, and of type $({}^{3}{}_{2}{}^{2}{}_{1}{}^{2})$. We now introduce the idea of a contraction which modifies the type of a tensor product.

Note/"Definition." A *contraction map* of a tensor product of copies of vector spaces X and X^* is a linear mapping to a tensor product involving one less copy of X and one less copy of X^*. This is denoted $\underset{\sim}{C}{}^i_j$ where we eliminate the ith contravariant space and the j covariant space from the original tensor product. For example, with

$$\underset{\sim}{C}{}^3_2 : X \otimes X \otimes \underset{\uparrow}{X} \otimes X^* \otimes X \otimes \underset{\uparrow}{X^*} \otimes X^* \to X \otimes X^* \otimes X \otimes X^*$$

we define $\underset{\sim}{C}{}^3_2$ on simple tensors as

$$\underset{\sim}{C}{}^3_2(\mathbf{x} \otimes \mathbf{x}_2 \otimes \mathbf{x}_2 \otimes \mathbf{f}_1 \otimes \mathbf{x}_4 \otimes \mathbf{f}_2 \otimes \mathbf{f}_3) = (\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{f}_1 \otimes \mathbf{x}_4 \otimes \mathbf{f}_3)(\mathbf{f}_2(\mathbf{x}_3));$$

notice that defining $\underset{\sim}{C}{}^3_2$ on simple tensors determines the contraction on the whole tensor product space since the simple tensors span the tensor product space (as in Exercise V.1.4(a)).

Note. We now dip into setting where we potentially have very numerous indices. When we introduce bases for vector space, this will let us express tensors in a tensor product space with respect to the resulting basis for the tensor product space. Instead of having an n-tuple representation (as in $\mathbb{R}^n$ or n-dimensional vector space X) we will have n^k components when the tensor product space results from a product of k n-dimensional spaces.

Note. If $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is a basis for vector space X and $\{\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^n\}$ is the dual basis for X^*, then by Exercise V.1.4(c)

$$\{\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \otimes \mathbf{b}^\ell \otimes \mathbf{b}^m \mid 1 \leq i, j, k, \ell, m \leq n\}$$

is a basis for $X \otimes X \otimes X \otimes X^* \otimes X^* = X_2^3$. So any $\mathbf{x} \in X_2^3$ we have $\mathbf{x}$ as a unique linear combination of the form

$$\mathbf{x} = (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \otimes \mathbf{b}^\ell \otimes \mathbf{b}^m) x_{\ell m}^{ijk}.$$

So we can represent tensor $\mathbf{x} \in X_2^3$ in terms of its components $x_{\ell m}^{ijk}$ (notice that this requires n^5 real numbers to thusly represent $\mathbf{x}$).

Note. Recall from Section III.1 (see Theorem III.1.B and the note that follows it) that to change the representation of a vector from basis β to basis β', we have the relationship between the coordinates of $x'^{\,i} = b_j^i x^j$ where $\mathbf{b}_i' = b_i^j \mathbf{b}_j$ and $\tilde{b}_k^i b_j^k = \delta_j^i$. If we convert tensor $\mathbf{x} \in X_2^3$ with coordinates $x_{\ell m}^{ijk}$ from basis

$$\{\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \otimes \mathbf{b}^\ell \otimes \mathbf{b}^m \mid 1 \leq i, j, k, \ell, m \leq n\}$$

to basis

$$\{\mathbf{b}'_i \otimes \mathbf{b}'_j \otimes \mathbf{b}'_k \otimes \mathbf{b}^{\ell\,\prime} \otimes \mathbf{b}^{m\,\prime} \mid 1 \le i, j, k, \ell, m \le n\}$$

we get "by precisely similar arguments" as given in Section III.1 that

$$x^{i'j'k'}_{\ell'm'} = \tilde{b}^{i'}_i \tilde{b}^{j'}_j \tilde{b}^{k'}_k b^{\ell}_{\ell'} b^{m}_{m'} x^{ijk}_{\ell m}$$

where $\mathbf{b}'_p = b^u_p \mathbf{b}_i$, $\mathbf{b}_i = \tilde{b}^p_i \mathbf{b}'_p$, and $\tilde{b}^{i'}_i = \delta^{i'}_p$. Similarly, a tensor of type $\binom{3\ \ 2\ \ 2}{\ \ 2\ \ 1\ \ }$ has coordinates labeled $x^{ijk\ \ m}_{\ \ \ \ \ell\ \ np}$ which transforms as

$$x^{i'j'k'\ \ m'}_{\ \ \ \ \ \ \ \ell'\ \ \ n'p'} = (\tilde{b}^{i'}_i \tilde{b}^{j'}_j \tilde{b}^{k'}_k b^{\ell'}_{\ell} b^{m'}_m b^{n'}_n b^{p'}_p) x^{ijk\ \ m}_{\ \ \ \ \ell\ \ np}.$$

Note. We can use coordinates and bases to represent tensor products of tensors. If

$$\mathbf{v} = v^{i_1 i_2 \cdots i_k}_{j_1 j_2 \cdots j_h} (\mathbf{b}_{i_1} \otimes \mathbf{b}_{i_2} \otimes \cdots \otimes \mathbf{b}_{i_k} \otimes \mathbf{b}^{j_1} \otimes \mathbf{b}^{j_2} \otimes \cdots \otimes \mathbf{b}^{j_n}) \in X^k_h$$

and

$$\mathbf{w} = w^{a_1 a_2 \cdots a_\ell}_{b_1 b_2 \cdots b_n} (\mathbf{b}_{a_1} \otimes \mathbf{b}_{a_2} \otimes \cdots \otimes \mathbf{b}_{a_\ell} \otimes \mathbf{b}^{b_1} \otimes \mathbf{b}^{b_2} \otimes \cdots \otimes \mathbf{b}^{b_m}) \in X^\ell_m$$

then

$$\mathbf{v} \otimes \mathbf{w} = v^{i_1 i_2 \cdots i_k}_{j_1 j_2 \cdots j_h} w^{a_1 a_2 \cdots a_\ell}_{b_1 b_2 \cdots b_n} (\mathbf{b}_{i_1} \otimes \mathbf{b}_{i_2} \otimes \cdots \otimes \mathbf{b}_{i_k} \otimes \mathbf{b}^{j_1} \otimes \mathbf{b}^{j_2} \otimes \cdots \otimes \mathbf{b}^{j_n})$$

$$\otimes (\mathbf{b}_{a_1} \otimes \mathbf{b}_{a_2} \otimes \cdots \otimes \mathbf{b}_{a_\ell} \otimes \mathbf{b}^{b_1} \otimes \mathbf{b}^{b_2} \otimes \cdots \otimes \mathbf{b}^{b_m}) \in X^{k\ \ \ell}_{\ \ h\ \ m} \cong X^{k+h\ell}_{h+m}$$

where the isomorphism holds by Exercise V.1.8(b).

Note. As an example of how a contraction affects coordinates, consider $\mathbf{x} = (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}^\ell \otimes \mathbf{b}^m) x^{ijk}_{\ell m} \in X^3_2$ and the contraction $\mathcal{C}^2_1$. Notice

$$\mathcal{C}^2_1(\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \otimes \mathbf{b}^\ell \otimes \mathbf{b}^m = \mathbf{b}_i \otimes \mathbf{b}_k \otimes \mathbf{b}^m(\mathbf{b}^\ell(\mathbf{b}_j)) = (\mathbf{b}_i \otimes \mathbf{b}_k \otimes \mathbf{b}^m)\delta^i_j$$

since, by the definition of dual basis, $\mathbf{b}^\ell(\mathbf{b}_j) = \delta^\ell_j$, so

$$\underset{\sim}{C}^2_1(\mathbf{x}) = (\mathbf{b}_i \otimes \mathbf{b}_k \otimes \mathbf{b}^m)\delta^\ell_j x^{ijk}_{\ell m} = (\mathbf{b}_i \otimes \mathbf{b}_k \otimes \mathbf{b}^m)x^{ijk}_{jm} \in X^2_1.$$

Notice that, by convention, there is summation over parameter j in the scalars so that neither j nor ℓ explicitly appear in the tensor sums; this is called *contracting over j and ℓ.*

Note. Recall that in the proof of Lemma V.1.08, we had $\mathbf{f} : X_1^* \otimes X_2 \to L(X_1; X_2)$ as $\mathbf{f}((\mathbf{g}, \mathbf{x}_2)) = \mathbf{h}$ where $\mathbf{h}(\mathbf{x}_1) = \mathbf{x}_2(\mathbf{g}(\mathbf{x}_1))$, and isomorphism $\hat{\mathbf{f}} : X_1^* \otimes X_2 \to L(X_1; X_2)$ where $\mathbf{f} = \hat{\mathbf{f}} \circ \bigotimes$. Let $\mathbf{a} \in X_1^* \otimes X_2$ where $\mathbf{a} = \mathbf{b}^j \otimes \mathbf{b}'_i a^i_j$ for bases $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ of X_1 and $\{\mathbf{b}'_1, \mathbf{b}'_2, \ldots, \mathbf{b}'_m\}$ of X_2 (so that the dual basis for X_1^* is $\{\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^n\}$). Then for $\mathbf{x} \in X_1$ we have

$$\begin{aligned}
(\hat{\mathbf{f}}\mathbf{a})\mathbf{x} &= \hat{\mathbf{f}}(\mathbf{b}^i \otimes \mathbf{b}'_i a^i_j)\mathbf{x} \\
&= \hat{\mathbf{f}}(\mathbf{b}^j \otimes \mathbf{b}'_i)\mathbf{x} a^i_j \text{ since } \hat{\mathbf{f}} \text{ is linear as shown in the proof of Lemma V.1.08} \\
&= \hat{\mathbf{f}}\left(\bigotimes((\mathbf{b}^k, \mathbf{b} - i'))\right)\mathbf{x} a^i_j \text{ by the definition of } \bigotimes \\
&= \mathbf{f}((\mathbf{b}^k, \mathbf{b}'_i))\mathbf{x} a + j^i \text{ since } \mathbf{f} = \hat{\mathbf{f}} \circ \bigotimes \\
&= \mathbf{b}'_i(\mathbf{b}^j(\mathbf{x}))a^i_j \text{ by the definition of } \mathbf{f} \\
&= \mathbf{b}'_i(\mathbf{b}^j(x^\ell \mathbf{b}_\ell))a^i_j = (\mathbf{b}'_i x^j)a^i_j \text{ by the definition of dual basis } \mathbf{b}^j.
\end{aligned}$$

So the matrix representing $\hat{\mathbf{f}}\mathbf{a} \in L(X_1; X_2)$ is the same as the matrix representing $\mathbf{a} \in X_1^* \otimes X_2$, namely $[\mathbf{a}] = [a^i_j]$. So whatever bases we choose for X_1 and X_2, they give the same representation for both $\hat{\mathbf{f}}\mathbf{a}$ and $\mathbf{a}$.

Note. In Theorem IV.1.09 and Note IV.1.A, we saw that in metric vector space $(X, \mathbf{G})$, $\mathbf{G}_\downarrow : X \to X^*$ and $\mathbf{G}_\uparrow : X^* \to X$ are isomorphisms. These yield isomorphisms between certain tensor products of spaces. For example,

$$\mathbf{\Theta} = \mathbf{I} \otimes \mathbf{I} \otimes \mathbf{G}_\uparrow \otimes \mathbf{G}_\downarrow \otimes \mathbf{G}_\uparrow \otimes \mathbf{I} \otimes \mathbf{G}_\uparrow : X \times X^* \otimes X^* \otimes X \otimes X^* \otimes X \otimes X^*$$

$$\to X \otimes X^* \otimes X \otimes X^* \otimes X \otimes X \otimes X$$

is an isomorphism. More generally, we have $X_h^k \cong X_{h'}^{k'}$ if $k + h = k' + h'$; this is unimpressive since both are real vector spaces with, by Exercise V.1.4(c), the same dimension. Additionally, it seems desirable to keep the original structure of a tensor product of spaces. Dodson and Poston state "velocity at a point arises as a contravariant vector. The gradient of a potential at a point arises as a covariant one and the contours of the functional ... are the local linear approximation to those of the potential." See page 109.

Note. Let $\mathbf{A} : X \to Y$ and $\mathbf{A}' : X' \to Y'$ have matrix representations $[a_j^i]$ and $[a'^k_\ell]$ with respect to bases $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, $\{\mathbf{b}'_1, \mathbf{b}'_2, \ldots, \mathbf{b}'_n\}$, $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m\}$, and $\{\mathbf{c}'_1, \mathbf{c}'_2, \ldots, \mathbf{c}'_m\}$ for X, X', Y, and Y', respectively. Then for $\mathbf{x} \in X \otimes X'$ we have

$$\begin{aligned}
\mathbf{A} \otimes \mathbf{A}'(\mathbf{x}) &= \mathbf{A} \otimes \mathbf{A}'(\mathbf{b}_j \otimes \mathbf{b}'_i x^{j\ell}) \text{ since } \{\mathbf{b}_j \otimes \mathbf{b}'_i \mid 1 \le j, \ell \le n\} \text{ is a basis for} \\
&\quad X \otimes X' \text{ by Exercise V.1.4(c)} \\
&= \mathbf{A} \otimes \mathbf{A}'(\mathbf{b}_j \otimes \mathbf{b}'_\ell) x^{j\ell} \\
&= (\mathbf{A}\mathbf{b}_j \otimes \mathbf{A}'\mathbf{b}_{\ell'}) x^{j\ell} \text{ by the definition of the tensor product of maps} \\
&= (\mathbf{c}_i a_j^i) \otimes (\mathbf{c}'_k a'^k_\ell) x^{j\ell} \text{ since } \mathbf{A}\mathbf{b}_j \in Y,\ \mathbf{A}'\mathbf{b}_\ell \in Y',\ \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m\} \text{ is a} \\
&\quad \text{basis of } Y,\ \{\mathbf{c}'_1, \mathbf{c}'_2, \ldots, \mathbf{c}'_m\} \text{ is a basis for } Y',\ [a_j^i] \text{ represents } \mathbf{A},
\end{aligned}$$

$$\begin{aligned} & \text{and } [a'^k_\ell] \text{ represents } \mathbf{A}' \\ = & \ (\mathbf{c}_i \otimes \mathbf{c}'_k) a^i_j a'^k_\ell x^{j\ell}. \end{aligned}$$

So the matrix representation of $\mathbf{A} \otimes \mathbf{A}'$ is the $nn' \times mm'$ matrix with entries $a^i_j a'^k_\ell$ for $1 \le i \le m$, $1 \le j \le n$, $1 \le k \le m'$, and $1 \le \ell \le n'$. The the mapping $\boldsymbol{\Theta} = \mathbf{I} \otimes \mathbf{I} \otimes \mathbf{G}_\uparrow \otimes \mathbf{G}_\downarrow \otimes \mathbf{G}_\uparrow \otimes \mathbf{I} \otimes \mathbf{G}_\uparrow$ can be represented by a matrix with entries which are the products of all elements representing the constituents of the tensor product of maps in $\boldsymbol{\Theta}$. So for $\boldsymbol{\Theta}(\mathbf{x}) = \mathbf{y}$ we have the transformation of coordinates

$$y^{i\ k'}_{\ j\ \ \ell'}{}^{m'np'} = g^{k'k} g_{\ell'\ell} g^{m'm} g^{p'p} x^i{}_{jk}{}^{\ell}{}_m{}^n{}_p.$$

Notice that in terms of **coordinates**, $\mathbf{G}_\uparrow$ "raises" indices k, m, and p, and $\mathbf{G}_\downarrow$ "lowers" index ℓ. *This* is why we use the up-arrow and down-arrow notation!

Note. Consider $\hat{\mathbf{f}} : X^* \otimes X \to L(X;X)$ from the proof of Lemma V.1.08, and

$$\mathbf{I} \otimes \mathbf{G}_\downarrow : X^* \otimes X \to X^* \otimes X^* = L(X, X; \mathbb{R}) = L^2(X; \mathbb{R})$$

(recall that $X^* \otimes X^*$ is defined to be $L(X, X; \mathbb{R})$). Define $\boldsymbol{\Psi} : L(X;X) \to L^2(X;\mathbb{R})$ as $\boldsymbol{\Psi} = \mathbf{I} \times \mathbf{G}_\downarrow \circ \hat{\mathbf{f}}^{-1}$. Then $\boldsymbol{\Psi}$ is an isomorphism between the space of linear operators $L(X;X)$ and the space of bilinear forms $L^2(X;\mathbb{R})$. More directly, for $\mathbf{A} \in L(X;X)$ we have $\boldsymbol{\Psi}(\mathbf{A})((\mathbf{x}, \mathbf{y})) = \mathbf{A}\mathbf{x} \cdot \mathbf{y}$ by Exercise V.1.10(a). Additionally, $\mathbf{A}$ is nonsingular if and only if $\boldsymbol{\Psi}\mathbf{A}$ is non-degenerate, and $\mathbf{A}$ is self-adjoint if and only if $\boldsymbol{\Psi}\mathbf{A}$ is symmetric (by Exercise V.1.10(b)).

Revised: 5/25/2019