Introduction to Tensors[1]

To understand tensors, one needs to understand three kinds of vector spaces and be able to switch back and forth between them: abstract vector spaces, vector spaces of linear functionals, and vector spaces of matrices/column vectors/row vectors. This is because tensors are defined as linear functionals that are defined over abstract vector spaces, and computations are often carried out using matrix algebra.

For this reason, we begin by providing brief descriptions of abstract vector spaces, linear functionals, and matrices. We use coordinate-free definitions, to introduce vector-space objects, which is the preferred approach. Then coordinate systems are introduced in order to define covectors spaces, which are involved in the definition of tensor.

[1] This introduction is based on Appendix C of "A Short Course in General Relativity" by Foster and Nightingale, 3rd edition, Springer Science+Business Media, 2006, and "Topics in algebra" by Herstein, Blaisdell Publishing Co., 1964.

1 Vector spaces

1A Abstract Vector Spaces

Definition **An *n*-dimensional vector space $\mathcal{V}$ over a field F** is an abstract collection of objects called **vectors** that is closed under addition and scalar multiplication:

$a\boldsymbol{V} + b\boldsymbol{W} \in \mathcal{V}$ if a and b are scalars and $\boldsymbol{V}$ and $\boldsymbol{W}$ are vectors

For quantum mechanics the field **F** is the complex numbers, $\mathbb{C}$. For relativity theory it is the reals, $\mathbb{R}$. A complete vector space definition includes several rules that ensure proper behavior of addition and scalar multiplication such as commutative and associative properties, which will not be listed here.

Definition **Scalars** are simply elements of **F**.

Theorem Every *n*-dimensional vector space has a **basis**, a linearly independent set of vectors that span the vector space. That is, letting $\mathcal{B}_{\mathcal{V}} = \left\{ \mathbf{e}_{(\mu)} \right\}$ denote a basis for $\mathcal{V}$, every vector in $\mathcal{V}$ can be uniquely expressed as a **linear combination** of basis vectors:

$$\boldsymbol{V} = V^{\mu}\mathbf{e}_{(\mu)} = V^{1}\mathbf{e}_{(1)} + \cdots + V^{n}\mathbf{e}_{(n)} .$$

The expression above for $\boldsymbol{V}$ uses the Einstein summation convention to add products of terms having matching upper and lower Greek indices. Parentheses on the subscripts emphasize that the objects consist of a set vectors, not a set of components of a single vector. The lack of parentheses on the superscripts of V emphasizes that they are the components. Vectors are bolded. Components, which are scalars, will not be put in bold face.

Definitions When every vector can be expressed as above by a linear combination of basis vectors, we say that $\mathcal{B}_{\mathcal{V}}$ **spans** $\mathcal{V}$. A set of vectors $\left\{ \mathbf{e}_{(\mu)} \right\}$ is **linearly independent** means that if there are scalars α^{μ} such that $\alpha^{\mu}\mathbf{e}_{(\mu)} = 0$ then $\alpha^{\mu} = 0 \;\; \forall \mu$.

Note that having a basis is paramount to having a coordinate system. That is, a vector $\boldsymbol{V}$ expressed in terms of a basis could equivalently be considered to be a point $\boldsymbol{V}$ having coordinates: $\boldsymbol{V} = (V^1, \ldots V^n)$. Relationships that depend on bases are therefore coordinate-dependent relationships.

Note, also, that the existence of a basis requires a theorem (not proven here). The process of finding a basis can vary greatly for different vector spaces.

Definition Vector spaces $\mathcal{V}$ and $\mathcal{W}$ are **isomorphic** means there is a 1-1 and onto mapping $T : \mathcal{V} \to \mathcal{W}$ that preserves the vector space operations of addition and scalar multiplication:

$$T(\alpha U + \beta V) = \alpha T(U) + \beta T(V),$$

where the '+' on the left occurs in $\mathcal{V}$ and the '+' on the right occurs in $\mathcal{V}$. A map T that preserves the vector space operations is said to be a **linear map**.

Theorem All n-dimensional vector spaces are isomorphic.
Proof: Let $\mathcal{V}$ and $\mathcal{W}$ be n-dimensional vector spaces. Then they have bases $\{\boldsymbol{e}_i : I = 1, \ldots, n\}$ and $\{\boldsymbol{f}_i : I = 1, \ldots, n\}$, respectively. The linear map T that takes $\boldsymbol{e}_i = \boldsymbol{f}_i$ is the required isomorphism. ■

1B Linear Functionals

An important class of vector spaces are vector spaces whose elements are functions.

Definition Let $X = \{x\}$ and $Y = \{y\}$ be sets. We say that $f : X \to Y : f(x) = y$ is a **function** if f assigns a unique y to every x.

Theorem The collection of functions $\{ f \}$ from X to Y form a vector space when addition and scalar multiplication of functions are defined naturally:

1. $(f+g)(x) = f(x) + g(x)$
2. $(\alpha f)(x) = \alpha[f(x)]$

Proof of this theorem is straightforward and easy but is omitted, as we have not listed the rules that define a vector space. This theorem is found in every book on linear algebra.

Definition Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces. A **linear map** $T : \mathcal{V} \to \mathcal{W}$ is a function that preserves the vector space structure (of addition and scalar multiplication):

1. $T(V+W) = T(V) + T(W)$
2. $T(\alpha V) = \alpha T(V)$

Definition A linear map $T : \mathcal{V} \to \mathbf{F}$ is called a **linear functional over $\mathcal{V}$.** The collection of all linear functionals over $\mathcal{V}$ is called the **dual space** of $\mathcal{V}$ and is denoted $\mathcal{V}^*$. The elements T of $\mathcal{V}^*$ are called **dual vectors**.

Theorem $\mathcal{V}^*$ is an n-dimensional vector space with addition and scalar multiplication defined naturally:

1. $(\alpha T)(V) = \alpha[T(V)]$
2. $(T_1 + T_2)(V) = T_1(V) + T_2(V)$

Proof: $\mathcal{V}^*$ is a vector space of functions by theorem above. ■

Definition $\mathcal{V}^{**}$ denotes the **second dual space** of an abstract vector space $\mathcal{V}$. It is the dual space of the dual space**.**

Theorem $\mathcal{V}^{**}$ is isomorphic to $\mathcal{V}$.

Proof: For $\boldsymbol{V} \in \mathcal{V}$ define $\mathfrak{T}_{\boldsymbol{V}} : \mathcal{V}^* \to \mathbf{F} : \mathfrak{T}_{\boldsymbol{V}}(T) = T(\boldsymbol{V})$. It is straightforward to show that $\mathfrak{T}_{\boldsymbol{V}}(\alpha T_1 + \beta T_2) = \alpha \mathfrak{T}_{\boldsymbol{V}}(T_1) + \beta \mathfrak{T}_{\boldsymbol{V}}(T_2)$. Therefore $\mathfrak{T}_{\boldsymbol{V}} \in \mathcal{V}^{**}$.

Define $h : \mathcal{V} \to \mathcal{V}^{**} : h(\boldsymbol{V}) = \mathfrak{T}_{\boldsymbol{V}}$. Again, it is straightforward to show that $h(\alpha \boldsymbol{V} + \beta \boldsymbol{W}) = \alpha h(\boldsymbol{V}) + \beta h(\boldsymbol{W})$. So, h is a linear map from $\mathcal{V}$ to $\mathcal{V}^{**}$. It is easy to show that h is 1-1 and onto. Thus, h is an isomorphism of $\mathcal{V}$ onto $\mathcal{V}^{**}$. ■

The importance of this theorem is that a vector in $\mathcal{V}$ can be treated as a linear functional over $\mathcal{V}^*$. Since, also, a vector in $\mathcal{V}^*$ is a linear functional over $\mathcal{V}$,

> $\mathcal{V}$ and $\mathcal{V}^*$ can be considered to be dual spaces of each other.

$\mathcal{V}^*$ has a basis since it is a vector space. Bases are not unique. We are interested in a special basis that is related to the basis $\mathcal{B}_{\mathcal{V}} = \left\{\mathbf{e}_{(\mu)}\right\}$ of $\mathcal{V}$.

Construction We begin by defining a functional only on $\mathcal{B}_{\mathcal{V}}$. Define

$$\boxed{e^{(\nu)} : \mathcal{B}_{\mathcal{V}} \to \mathbf{F} : e^{(\nu)}\left(\boldsymbol{e}_{(\mu)}\right) = \delta^{\nu}_{\mu}}. \tag{1.1}$$

Next, we extend $e^{(\nu)}$ linearly to all of $\mathcal{V}$:

$$e^{(\nu)} : \mathcal{V} \to \mathbf{F} : e^{(\nu)}(\boldsymbol{V}) = e^{(\nu)}\left(V^{\mu} \mathbf{e}_{\mu}\right) = V^{\mu} e^{(\nu)}\left(\mathbf{e}_{\mu}\right) = V^{\mu} \delta^{\nu}_{\mu} = V^{\nu}.$$

Claim $e^{(\nu)}$ is a linear functional:

$$e^{(\nu)}\left(a\boldsymbol{V}+b\boldsymbol{W}\right)=e^{(\nu)}\left(aV^{\mu}\boldsymbol{e}_{\mu}+bW^{\mu}\boldsymbol{e}_{\mu}\right)=e^{(\nu)}\left[\left(aV^{\mu}+bW^{\mu}\right)\boldsymbol{e}_{\mu}\right]$$

$$\overset{\text{(linear)}}{=}\left(aV^{\mu}+bW^{\mu}\right)e^{(\nu)}\left(\boldsymbol{e}_{\mu}\right)=\left(aV^{\mu}+bW^{\mu}\right)\delta_{\mu}^{\nu}$$

$$=aV^{\nu}+bW^{\nu}=a\,e^{(\nu)}\left(\boldsymbol{V}\right)+b\,e^{(\nu)}\left(\boldsymbol{W}\right) \qquad ✔$$

Finally, claim $\mathcal{B}_{\mathcal{V}}{}^{*}\equiv\left\{e^{(\nu)}\right\}_{\nu=1}^{n}$ is a basis for $\mathcal{V}^{*}$.

We must show that $\{e^{(\nu)}\}$ spans $\mathcal{V}^{*}$ and is linearly independent.

<u>Span</u>: Let $T\in\mathcal{V}^{*}$. Set $T_{\mu}=T\left(\boldsymbol{e}_{(\mu)}\right)\in\mathbf{F}$. Then $T=T_{\nu}e^{(\nu)}$ because for all $\boldsymbol{V}$

$$T\left(\boldsymbol{V}\right)=V^{\mu}T\left(\boldsymbol{e}_{(\mu)}\right)=T_{\mu}V^{\mu}=T_{\nu}V^{\mu}\delta_{\mu}^{\nu}=T_{\nu}V^{\mu}e^{(\nu)}\left(\boldsymbol{e}_{(\mu)}\right)=T_{\nu}e^{(\nu)}\left(V^{\mu}\boldsymbol{e}_{(\mu)}\right)=T_{\nu}e^{(\nu)}\left(\boldsymbol{V}\right) ✔$$

<u>Linearly independent</u>: Suppose scalars α_{ν} exist such that $\alpha_{\nu}e^{(\nu)}=0$. Then

$$0=\alpha_{\nu}e^{(\nu)}=\alpha_{\nu}e^{(\nu)}\left(\boldsymbol{e}_{(\mu)}\right)=\alpha_{\nu}\delta_{\mu}^{\nu}=\alpha_{\mu}\quad\forall\mu \qquad ✔ \qquad ■$$

Since the basis $\mathcal{B}_{\mathcal{V}}{}^{*}$ has *n* members, we have discovered the following in passing.

Corollary The dimension of $\mathcal{V}^{*}$ is the same as the dimension of $\mathcal{V}$.

Definition When a dual space is equipped with the specific basis $\mathcal{B}_{\mathcal{V}}{}^{*}$, we call $\mathcal{V}^{*}$ a **covector space** and the linear functionals are called **covectors**.

Covectors and vectors are on equal footing: each is a vector space and each is a function space, the dual of the other. When we treat their members as vectors, the symbols will be bolded: $\boldsymbol{V}$, $\boldsymbol{e}_{(\mu)}$, $\boldsymbol{\omega}$, and $\boldsymbol{e}^{(\nu)}$. When we treat them as functions, the symbols will not bolded: V, $e_{(\mu)}$, ω, and $e^{(\nu)}$.

The analog of equation (1.1) for elements $e_{(\mu)}$ of $\mathcal{V}$ is therefore written

$$e_{(\rho)}:\mathcal{B}_{\mathcal{V}^{*}}\to\mathbf{F}:e_{(\rho)}\left(\boldsymbol{e}^{(\sigma)}\right)=\delta_{\rho}^{\sigma}. \tag{1.2}$$

Remark Many books use the terms "covector", "linear functional", "dual vector", and "1-form" interchangeably. There are actually subtle differences between them. For the record, we distinguish them now.

1. When the term **linear functional** is used, the object is being considered as a function.
2. When the set of linear functionals is considered as a vector space, the vector space is known as the **dual space** and the linear functionals are referred to as **dual vectors.**
3. A **covector space** is a dual space characterized by having basis $\left\{e^{(\nu)}\right\}$ where $e^{(\nu)}\left(\mathbf{e}_{(\mu)}\right)=\delta_{\mu}^{\nu}$. The vectors of a covector space are called **covectors**.
4. The covector basis elements, $\mathbf{e}^{(\nu)}$, can be constructed as vectors that are perpendicular to the basis elements $\mathbf{e}_{(\mu)}$. In this construction they can be expressed as differentials, ∂f, and called **1-forms**.

1C <u>Matrices, Column Vectors, and Row Vectors</u>

Notation **Column vectors** will be denoted in bold upper case by $\boldsymbol{C}$ and $\boldsymbol{D}$. **Row vectors** will be denoted in bold lower case by $\boldsymbol{r}=\begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix}$ and $\boldsymbol{s}=\begin{pmatrix} s_1 & \cdots & s_n \end{pmatrix}$. If $\boldsymbol{C}$ is a column vector, then its **transpose**, $\boldsymbol{C}^{\mathrm{T}}$, is a row vector. The symbol $\mathcal{C}$ is used to denote the vector space of length-*n* column vectors. $\mathcal{R}$ is used to denote the vector space of length-*n* row vectors.

Theorem The collections $\mathcal{C}$ and $\mathcal{R}$ are *n*-dimensional vector spaces with addition and scalar multiplication defined naturally:

$$\boldsymbol{C}+\boldsymbol{D}\equiv\begin{pmatrix} V^1 \\ \vdots \\ V^\mu \\ \vdots \\ V^n \end{pmatrix}+\begin{pmatrix} W^1 \\ \vdots \\ W^\mu \\ \vdots \\ W^n \end{pmatrix}=\begin{pmatrix} V^1+W^1 \\ \vdots \\ V^\mu+W^\mu \\ \vdots \\ V^n+W^n \end{pmatrix} \quad \text{and} \quad \alpha\boldsymbol{C}=\begin{pmatrix} \alpha V^1 \\ \vdots \\ \alpha V^\mu \\ \vdots \\ \alpha V^n \end{pmatrix}$$

$$\boldsymbol{r}+\boldsymbol{s}=\left(r_1\cdots r_n\right)+\left(s_1\cdots s_n\right)=\left(r_1+s_1\cdots r_n+s_n\right) \quad \text{and} \quad \alpha\boldsymbol{r}=\left(\alpha r_1\cdots\alpha r_n\right)$$

Notation Multiple column vectors like $\boldsymbol{C}_{(1)}$ and $\boldsymbol{C}_{(2)}$ will be denoted in bold upper case. Subscripts are put in parentheses to emphasize they are vectors rather than components of vectors. Basis column vectors, like $\boldsymbol{E}_{(\mu)}$, will be denoted in bold lower case with subscripts put in parentheses. Row vectors and basis row vectors will be denoted $\boldsymbol{r}^{(\nu)}$ and $\boldsymbol{E}^{(\nu)}$, respectively.

Construction 1.1 Abstract vectors are identified with column vectors

To identify an abstract vector $\boldsymbol{V}$ with a column vector $\boldsymbol{C_V}$ we first generate a basis $\mathcal{B}=\left\{\mathbf{e}_{(\mu)}\right\}$ for $\mathcal{V}$. Next we identify $\mathcal{B}$ with the basis $\{\boldsymbol{E}_{(\mu)}\}$ of $\mathcal{C}$:

$$\mathbf{e}_{(\mu)} \leftrightarrow \boldsymbol{E}_{(\mu)} \equiv \begin{pmatrix} 0 \\ \vdots \\ 1^{\mu} \\ \vdots \\ 0 \end{pmatrix}.$$

Finally, we identify $\boldsymbol{V}=V^{\mu}\mathbf{e}_{(\mu)} \in \mathcal{V}$ with the column vector $\boldsymbol{C_V}=V^{\mu}\boldsymbol{E}_{(\mu)} \in \mathcal{C}$:

$$\boldsymbol{V} \leftrightarrow \boldsymbol{C_V} \equiv \begin{pmatrix} V^1 \\ \vdots \\ V^{\mu} \\ \vdots \\ V^n \end{pmatrix}..$$

It is clear that the mapping $\boldsymbol{V} \mapsto \boldsymbol{C}$ is 1-1 and onto. It is also easy to see that addition and scalar multiplication in $\mathcal{C}$ matches that of $\mathcal{V}$. Thus, $\mathcal{V}$ is isomorphic with $\mathcal{C}$**.** This proves that every *n*-dimensional abstract vector space is isomorphic to the space of *n*-dimensional column vectors. ■

Construction 1.2 Abstract covectors are identified with row vectors

We use $\mathcal{B}_{\mathcal{V}^*}=\left\{\mathbf{e}^{(\nu)}\right\}$ as the basis for $\mathcal{V}^*$ and define the row vectors $\boldsymbol{E}^{(\nu)} \equiv \begin{pmatrix} 0 & \cdots & 1_{\nu} & \cdots & 0 \end{pmatrix}$ as the basis for $\mathcal{R}$. We identify the abstract covector $\boldsymbol{\omega}=\omega_{\nu}\mathbf{e}^{(\nu)}$ with the row vector $\boldsymbol{r}^{(\omega)}=r_{\nu}\boldsymbol{E}^{(\nu)}=\left(r_1 \cdots r_n\right)$. The map $\boldsymbol{\omega} \mapsto \boldsymbol{r}^{(\omega)}$ is an isomorphism between $\mathcal{V}^*$ and $\mathcal{R}$. ■

Row vectors are covectors to the column vectors because the equivalent of equation (1.1) is satisfied:

$$\boldsymbol{E}^{(\nu)}\boldsymbol{E}_{(\mu)} = \begin{pmatrix} \cdots & 1_\nu & \cdots \end{pmatrix}\begin{pmatrix} \vdots \\ 1^\mu \\ \vdots \end{pmatrix} = \delta^\nu_\mu$$

Theorem The product of any row vector with any column vector is a scalar:

$$\boldsymbol{r}\,\boldsymbol{C} = \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix}\begin{pmatrix} V^1 \\ \vdots \\ V^n \end{pmatrix} = r_\mu V^\mu \in \mathbf{F}. \tag{1.3}$$

Definition The **product $\mathcal{V}\times\mathcal{W}$ of two abstract vector spaces** is defined to be the vector space of pairs ($\boldsymbol{V}$, $\boldsymbol{W}$) where $\boldsymbol{V}\in\mathcal{V}$ and $\boldsymbol{W}\in\mathcal{W}$. Addition and scalar multiplication are defined naturally:

$$\left(\boldsymbol{V},\boldsymbol{W}\right)+\left(\boldsymbol{X},\boldsymbol{Y}\right)=\left(\boldsymbol{V}+\boldsymbol{X},\boldsymbol{W}+\boldsymbol{Y}\right) \text{ and } \alpha\left(\boldsymbol{V},\boldsymbol{W}\right)=\left(\alpha\boldsymbol{V},\alpha\boldsymbol{W}\right).$$

A basis for $\mathcal{V}\times\mathcal{W}$ can be constructed from the bases of $\mathcal{V}$ and $\mathcal{W}$ as follows. If $\left\{\boldsymbol{e}_{(\mu)}\right\}$ is a basis for $\mathcal{V}$ and $\left\{\boldsymbol{f}_{(\nu)}\right\}$ is a basis for $\mathcal{W}$, then it is straightforward to show that $\left\{\left(\boldsymbol{e}_{(\mu)},\boldsymbol{f}_{(\nu)}\right)\right\}$ is a basis for $\mathcal{V}\times\mathcal{W}$. Thus, if $\boldsymbol{V}=V^\mu\boldsymbol{e}_{(\mu)}$ and $\boldsymbol{W}=W^\nu\boldsymbol{f}_{(\nu)}$, then ($\boldsymbol{V}$, $\boldsymbol{W}$) can be expressed in terms of basis pairs as

$$\left(\boldsymbol{V},\boldsymbol{W}\right)=\left(V^\mu\boldsymbol{e}_{(\mu)},W^\nu\boldsymbol{f}_{(\nu)}\right)=V^\mu W^\nu\left(\boldsymbol{e}_{(\mu)},\boldsymbol{f}_{(\nu)}\right).$$

This product vector space definition also applies to abstract covector spaces. For example, let $\mathcal{V}^*$ be the covector space of $\mathcal{V}$, $\left\{e^{(\nu)}\right\}$ the basis for $\mathcal{V}^*$, and $\boldsymbol{\xi}\in\mathcal{V}^*$. Then vectors in $\mathcal{V}^*\times\mathcal{W}$ can be expressed as

$$\left(\boldsymbol{\xi},\boldsymbol{W}\right)=\xi_\mu W^\nu\left(\boldsymbol{e}^{(\nu)},\boldsymbol{f}_{(\nu)}\right).$$

2 Tensors

We now have the machinery to define tensors and tensor products. Though tensors can be defined for infinite-dimensional vector spaces, we will restrict our scope to finite-dimensional vector spaces.

Definition Let $\mathcal{V}$, …, $\mathcal{W}$ be finite-dimensional vector spaces, and Υ^*, …, $\mathcal{Z}^*$ be finite-dimensional covector spaces. Let k and ℓ be non-negative integers. A **rank** $(\boldsymbol{k},\ell)$ **tensor** is a multilinear map T from a product of covector spaces and vector spaces to **F**:

$$\boxed{T:\underbrace{\Upsilon^*\times\cdots\times\mathcal{Z}^*}_{k\text{ terms}}\times\underbrace{\mathcal{V}\times\cdots\times\mathcal{W}}_{\ell\text{ terms}}\to\mathbf{F}}. \tag{2.1}$$

Multilinear means the tensor T acts linearly in each of its arguments. For example, for a (1,1) tensor,

$$\begin{aligned}&T\left(a_1\omega^{(1)}+a_2\omega^{(2)},b_1\boldsymbol{V}_{(1)}+b_{(2)}\boldsymbol{V}_{(2)}\right)\\&=a_1b_1T\left(\omega^{(1)},\boldsymbol{V}_{(1)}\right)+a_1b_2T\left(\omega^{(1)},\boldsymbol{V}_{(2)}\right)+a_2b_1T\left(\omega^{(2)},\boldsymbol{V}_{(1)}\right)+a_2b_2T\left(\omega^{(2)},\boldsymbol{V}_{(2)}\right).\end{aligned}$$

The upper indices are called **contravariant indices** and the lower ones are called **covariant indices.**

"Multilinear" is neither stronger nor weaker than "linear". It is just different. By restricting linearity to each argument separately, this property respects that we never combine tensor product arguments. This will be emphasized in Section 3.

Addition and scalar multiplication of rank (k,ℓ) tensors S and T is defined by

$$\begin{aligned}&[aS+bT]\left(\omega^{(1)},\cdots,\omega^{(k)},V_{(1)},\cdots,V_{(\ell)}\right)\\&\equiv a\left[\,S\left(\omega^{(1)},\cdots,\omega^{(k)},V_{(1)},\cdots,V_{(\ell)}\right)\right]+b\left[\,T\left(\omega^{(1)},\cdots,\omega^{(k)},V_{(1)},\cdots,V_{(\ell)}\right)\right].\end{aligned} \tag{2.2}$$

Theorem The *set* of (k, ℓ) tensors forms a vector space.
Proof. Equation (2.2) is the definition given in Section 1B for forming a vector space of functions. ■

Most of the properties of tensors can be illustrated using just a single covector space with a single vector space. One such property is multiplication of tensors. We define the tensor product now for a simple case. The general definition is withheld until Section 6.

Definition Let $R:\mathcal{V}^*\rightarrow\mathrm{F}$ and $S:\mathcal{W}\rightarrow\mathbf{F}$ be rank (1, 0) and rank (0, 1) tensors, respectively. Define the **tensor product** $R\otimes S$ to be the tensor defined by

$$R\otimes S:\mathcal{V}^*\times\mathcal{W}:\boxed{R\otimes S(\xi,\boldsymbol{W})\equiv R(\xi)\,S(\boldsymbol{W})}. \tag{2.3}$$

An important property of $\otimes$ is that the arguments ξ and $\boldsymbol{W}$ are never combined in any way. They are kept separate. R operates only on $\mathcal{V}^*$ and S operates only on $\mathcal{W}$. In this sense the tensor product acts as a bookkeeper, keeping each argument in its own "bin".

Tensor computations are often performed by matrix multiplication. In Sections 5 and 6 we will show how to represent tensors by matrices. Prior to that it is helpful to learn how to manipulate matrices using the tensor product $\otimes$. This is the subject of Section 3.

3 Tensor Products of Matrices

Tensor products of matrices can be displayed in different ways. In this section we introduce a very common method for displaying the tensor product $\boldsymbol{A}\otimes\boldsymbol{B}$ where $\boldsymbol{A}$ and $\boldsymbol{B}$ are any matrix objects: scalars, row vectors, column vectors, or matrices. We first define '$\otimes$' for a scalar k and a matrix object $\boldsymbol{A}$, where $\boldsymbol{A}$ is either a matrix, a scalar, a row vector, or a column vector:

Definition

$$k\otimes\boldsymbol{A}=\boldsymbol{A}\otimes k\equiv k\boldsymbol{A} \tag{3.1}$$

Examples of (3.1) are:

$$k\otimes\ell=k\ell,\ k\otimes\begin{pmatrix} a & b \end{pmatrix}=\begin{pmatrix} ka & kb \end{pmatrix},\ k\otimes\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} ka \\ kb \end{pmatrix}, \text{ and}$$

$$k\otimes\begin{pmatrix} a & b \\ c & d \end{pmatrix}=k\begin{pmatrix} a & b \\ c & d \end{pmatrix}=\begin{pmatrix} ka & kb \\ kc & kd \end{pmatrix}$$

We next define '$\otimes$' for two matrix objects that are not scalars.

Definition

$$\boldsymbol{A}\otimes\boldsymbol{B}\equiv\begin{pmatrix} A_{11}\otimes\boldsymbol{B} & \cdots & A_{1n}\otimes\boldsymbol{B} \\ \vdots & & \vdots \\ A_{n1}\otimes\boldsymbol{B} & \cdots & A_{nn}\otimes\boldsymbol{B} \end{pmatrix} \tag{3.2}$$

This is a bootstrap definition, using (3.1), because $A_{11} - A_{nn}$ are scalars. We refer to this as a *pattern definition* since the operation is easily carried out by following the bootstrap pattern (3.1/ 3.2). We give examples.

Example 1: Tensor product of two matrices:

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{21} \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{21} \end{pmatrix}$$

$$\boldsymbol{A}\otimes\boldsymbol{B}\overset{(3.2)}{=}\begin{pmatrix} a_{11}\otimes\boldsymbol{B} & a_{12}\otimes\boldsymbol{B} \\ a_{21}\otimes\boldsymbol{B} & a_{21}\otimes\boldsymbol{B} \end{pmatrix}\overset{(3.1)}{=}\left(\begin{array}{cc|cc} a_{11}\otimes b_{11} & a_{11}\otimes b_{12} & a_{12}\otimes b_{11} & a_{12}\otimes b_{12} \\ a_{11}\otimes b_{21} & a_{11}\otimes b_{22} & a_{12}\otimes b_{21} & a_{12}\otimes b_{22} \\ \hline a_{21}\otimes b_{11} & a_{21}\otimes b_{12} & a_{22}\otimes b_{11} & a_{22}\otimes b_{12} \\ a_{21}\otimes b_{21} & a_{21}\otimes b_{22} & a_{22}\otimes b_{21} & a_{22}\otimes b_{22} \end{array}\right)$$

Example 2: Tensor product of a column vector and a row vector:

$$\boldsymbol{C}\otimes\boldsymbol{r} = \begin{pmatrix} C^1 \\ \vdots \\ C^n \end{pmatrix} \otimes \boldsymbol{r} \overset{(3.2)}{=} \begin{pmatrix} C^1 \otimes \boldsymbol{r} \\ \vdots \\ C^n \otimes \boldsymbol{r} \end{pmatrix} = \begin{pmatrix} C^1 \otimes \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} \\ \vdots \\ C^n \otimes \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} \end{pmatrix}$$

$$\overset{(3.1)}{=} \begin{pmatrix} C^1 \otimes r_1 & \cdots & C^1 \otimes r_n \\ \vdots & & \vdots \\ C^n \otimes r_1 & \cdots & C^n \otimes r_n \end{pmatrix} \text{ or } \begin{pmatrix} (C^1, r_1) & \cdots & (C^1, r_n) \\ \vdots & & \vdots \\ (C^n, r_1) & \cdots & (C^n, r_n) \end{pmatrix}$$

The tensor symbols of the ordered pairs inside the matrices are simply reminders not to combine the scalars C^μ and r_ν. A special case of this is the tensor product of a column basis vector and a row basis vector:

$$\boldsymbol{E}_\mu^\nu \equiv \boldsymbol{E}_{(\mu)} \otimes \boldsymbol{E}^{(\nu)} \equiv \begin{pmatrix} \vdots \\ 1^\mu \\ \vdots \end{pmatrix} \otimes \begin{pmatrix} \cdots & 1_\nu & \cdots \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1_\nu^\mu & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \tag{3.3}$$

1_ν^μ has two parts: $1_\nu^\mu = 1^\mu \otimes 1_\nu$. ■

Example 3: Tensor product of two row vectors:

$$\begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} \otimes \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} \overset{(3.2)}{=} \begin{pmatrix} r_1 \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} & \cdots & r_n \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} \end{pmatrix}$$

$$\overset{(3.1)}{=} \begin{pmatrix} r_1 r_1 & \cdots & r_1 r_n & r_2 r_1 & \cdots & r_2 r_n & \cdots & r_n r_1 & \cdots & r_n r_n \end{pmatrix}$$

As shown, as long as we don't combine $r_\nu r_\sigma$ it is acceptable to drop the tensor symbol. ■

We see that when $\boldsymbol{r}$ is a length n row vector, $\boldsymbol{r}\otimes\boldsymbol{r}$ is a length n^2 row vector. We see that when $\boldsymbol{C}$ has length n, $\boldsymbol{C}\otimes\boldsymbol{r}$ is an nxn matrix. Similarly, the tensor product of two length n column vectors would be expressed as a length n^2 column vector.

The pattern definition is easy to extend to triple tensor products and more. For a triple tensor product of row vectors, $\boldsymbol{r}^{(1)} \otimes \boldsymbol{r}^{(2)} \otimes \boldsymbol{r}^{(3)}$, the pattern definition would yield a 1xn^3 row vector.

We have chosen to represent the tensor product of two n-dimensional row vectors as an n^2-dimensional row vector, and this approach is often taken in Quantum Mechanics. In General Relativity it is often more convenient to represent this same tensor product as an nxn matrix. Because the vector spaces of nxn matrices, 1xn^2 row vectors, and n^2x1 column vectors have the same vector space dimension n^2, they are isomorphic to each other. So, representing the same tensor product as a row vector, column vector, or matrix is equally legitimate.

4 Lower Rank Tensors

We briefly summarize the lower rank tensors. The tensor definition does not cover rank (0, 0). A rank (0, 0) tensor is defined to be a scalar.

A rank (0, 1) tensor space is $\mathcal{T} = \{T : \mathcal{V} \to \mathbf{F} :\ T \text{ is linear}\}$. This is the definition of the dual space. When we impose the covector condition, equation (1.1), $\mathcal{T}$ becomes the covector space, $\mathcal{V}$*****. A rank (0, 1) tensor is a covector, and can be represented by the coefficient T_ν and identified with a row vector (by Construction 1.2).

A rank (1, 0) tensor space is $\mathcal{T} = \{T : \mathcal{V}^* \to \mathbf{F} :\ T \text{ is linear}\}$. Observe that this is the definition of the second dual space of $\mathcal{V}$. Thus, $\mathcal{T}$ is isomorphic to $\mathcal{V}$. A rank (1, 0) tensor is a vector, and can be represented by the coefficient T^μ and identified with a column vector (by Costruction 1.1).

A rank (1, 1) tensor space is $\mathcal{T} = \{T : \mathcal{V}^* \times \mathcal{W} \to \mathbf{F} :\ T \text{ is bilinear}\}$. This tensor space will be analyzed in depth in Section 5. It will be shown that a rank (1, 1) tensor can be represented by the coefficient T^μ_ν and identified with an nxm matrix. It will be shown ithat $\mathcal{T}$ contains tensor products $\xi \otimes W$ as a proper subset[2]. This means that, in general, we cannot decompose T^μ_ν into $T^\mu T_\nu$.

A rank (0, 2) tensor space is $\mathcal{T} = \{T : \mathcal{V} \times \mathcal{W} \to \mathbf{F} :\ T \text{ is bilinear}\}$.
Similar to rank (1, 1) tensors, this tensor space contains tensor products $\xi \otimes \omega$ as a proper subset. A rank (0, 2) tensor can be represented by the coefficient $T_{\nu\sigma}$ and identified with an nm row vector (as in Example 3 of Section 3).

A rank (2, 0) tensor space is $\mathcal{T} = \{T : \mathcal{V}^* \times \mathcal{W}^* \to \mathbf{F} :\ T \text{ is bilinear}\}$. This tensor space contains tensor products $V \otimes W$ as a proper subset. A rank (2, 0) tensor can be represented by the coefficient $T^{\mu\rho}$ and identified with an nm column vector.

---

[2] We show in Appendix A that this subset does not form a vector subspace and so, algebraically speaking, is not very interesting. However, in Quantum Mechanics, if ξ and W represent Alice's and Bob's spin states, the collection $\{\xi \otimes W\}$ is the set of spin product states and $\mathcal{T}$ is the set of entangled states. So, $\{\xi \otimes W\}$ is of interest, theoretically speaking.

5 Rank (1,1) Tensors

General tensors are functions of several covectors and several vectors. Many key facts about tensors can be illustrated using just one covector and one vector. We begin our investigation of tensor properties in this much simpler environment.

Tensors have been defined as functions. We generally think of tensors as being objects like $T^{\nu\rho}_{\mu\sigma}$ that have multiple subscripts and superscripts. Though we often refer to them as tensors, such objects are actually tensor coefficients and are tied to specific tensor bases. One objective of this section is to introduce tensor bases and coefficients and express tensors in terms of them. Another objective is to represent tensors in matrix form, useful for performing computations.

In equation (2.3) we gave a definition for the tensor product of a (1,0) tensor with a (0,1) tensor. Our last objective in this section is to extend that definition to the tensor product of two rank (1,1) tensors.

Below is an overview of the steps we will take in this section.

1. Tailor Definition (2.1) to a rank (1,1) tensor, T
2. Express the tensor T in terms of a tensor basis
3. Define the tensor product '$\otimes$' of a vector and a covector
4. Show that T can be represented by a matrix
5. Show how to express $T(\xi, \boldsymbol{W})$ using matrix algebra
6. Show that $T(\xi, \boldsymbol{W})$ can operate as a function of a single variable
7. Define the tensor product '$\otimes$' of two rank (1,1) tensors

<u>Step 1</u>: Tailor Definition (2.1) to a rank (1,1) tensor, *T*

Define a rank (1,1) tensor, T, as

$$T : \mathcal{V}^* \times \mathcal{W} \to \mathbf{F} : \ T \text{ is bilinear}. \tag{1}$$

The tensor space is the collection of all such T's:

$$\mathcal{T} = \{ T : T \text{ is a rank (1,1) tensor} \}.$$

<u>Step 2</u>: Express tensor *T* in terms of a tensor basis

We denote four basis sets, one each for $\mathcal{V}$, $\mathcal{W}$, $\mathcal{V}^*$, and $\mathcal{W}^*$.

$$\text{Vector space bases:} \quad \mathcal{B}_{\mathcal{V}} = \left\{ \mathbf{e}_{(\mu)} \right\}_{\mu=1}^{n}, \quad \mathcal{B}_{\mathcal{W}} = \left\{ \boldsymbol{f}_{(\rho)} \right\}_{\rho=1}^{m}$$

$$\text{Dual Space bases:} \quad \mathcal{B}_{\mathcal{V}^*} = \left\{ \mathbf{e}^{(\nu)} \right\}_{\nu=1}^{n}, \quad \mathcal{B}_{\mathcal{W}^*} = \left\{ \varphi^{(\sigma)} \right\}_{\sigma=1}^{m}. \tag{2}$$

Recall from equations (1.1) and (1.2) that covector bases satisfy

$$\mathbf{e}^{(\rho)}\left(\mathbf{e}_{(\mu)}\right)=\delta^{\rho}_{\mu},\quad \varphi^{(\nu)}\left(\boldsymbol{f}_{(\rho)}\right)=\delta^{\nu}_{\rho},\quad \mathrm{e}_{(\mu)}\left(\mathbf{e}^{(\sigma)}\right)=\delta^{\sigma}_{\mu},\ \text{and}\ f_{(\rho)}\left(\boldsymbol{\varphi}^{(\nu)}\right)=\delta^{\nu}_{\rho}. \tag{3}$$

We denote four vectors, one each for $\mathcal{V}$, $\mathcal{W}$, $\mathcal{V}^*$, and $\mathcal{W}^*$.

Vectors: $\boldsymbol{V}=V^{\mu}\,\mathbf{e}_{(\mu)}\in\mathcal{V},\quad \boldsymbol{W}=W^{\rho}\,\boldsymbol{f}_{(\rho)}\in\mathcal{W}$

Covectors: $\boldsymbol{\xi}=\xi_{\nu}\,\mathbf{e}^{(\nu)}\in\mathcal{V}^*,\quad \boldsymbol{\omega}=\omega_{\sigma}\,\boldsymbol{\varphi}^{(\sigma)}\in\mathcal{W}^*$ (4)

Remember: we use bold font for vectors and unbolded font for functions.

The usual choice of basis for $\mathcal{T}$ is $\left\{e^{\nu}_{\mu}\right\}$, defined by

$$e^{\nu}_{\mu}\left(\mathbf{e}^{(\sigma)},\ \boldsymbol{f}_{(\rho)}\right)\equiv\delta^{\sigma}_{\mu}\ \delta^{\nu}_{\rho} \tag{5}$$

The symbol, e^{ν}_{μ}, as defined so far, is simply a function that acts on basis pairs. To make it into a tensor, we extend it bilinearly to $\mathcal{V}^*\times\mathcal{W}$:

$$e^{\nu}_{\mu}\left(\xi_{\sigma}\mathbf{e}^{(\sigma)},W^{\rho}\boldsymbol{f}_{(\rho)}\right)=\xi_{\sigma}W^{\rho}e^{\nu}_{\mu}\left(\mathbf{e}^{(\sigma)},\boldsymbol{f}_{(\rho)}\right)$$

Claim $e^{\nu}_{\mu}\left(\boldsymbol{\xi},\boldsymbol{W}\right)=\xi_{\mu}W^{\nu}$: (6)

$$e^{\nu}_{\mu}\left(\boldsymbol{\xi},\boldsymbol{W}\right)=e^{\nu}_{\mu}\left(\xi_{\sigma}\mathbf{e}^{(\sigma)},W^{\rho}\boldsymbol{f}_{(\rho)}\right)\overset{\text{(bilinear)}}{=}\xi_{\sigma}W^{\rho}e^{\nu}_{\mu}\left(\mathbf{e}^{(\sigma)},\boldsymbol{f}_{(\rho)}\right)\overset{(5)}{=}\xi_{\sigma}W^{\rho}\delta^{\sigma}_{\mu}\ \delta^{\nu}_{\rho}=\xi_{\mu}W^{\nu}\ ✔$$

Sometimes there is a need to slant indices NW to SE or SW to NE:

1. Without slanting, one cannot immediately tell in (6) which index operates on $\boldsymbol{\xi}$ and which operates on $\boldsymbol{W}$. In fact, μ acts on $\boldsymbol{\xi}$ and ν acts on $\boldsymbol{W}$. When this is important, slanting $e_{\mu}{}^{\nu}$ makes it clear.

2. The metric tensor (not covered in this document) has an ability to raise and lower indices. Slanting is necessary to indicate, for example that the middle index is the one lowered in the following example:
$$e^{\alpha\beta\gamma}\rightarrow e^{\mu}{}_{\sigma}{}^{\nu}.$$

Denote the tensor basis for $\mathcal{T}$:

$$\mathcal{B}_{\mathcal{T}}=\left\{e^{\nu}_{\mu}\right\}^{n}_{\mu=1}{}^{m}_{\nu=1} \tag{7}$$

Claim: $\mathcal{B}_{\mathcal{T}}$ is a basis; i.e., $\mathcal{B}_{\mathcal{T}}$ spans $\mathcal{T}$ and is linearly independent:

$\mathcal{B}_{\mathcal{T}}$ spans $\mathcal{T}$: Let $T \in \mathcal{T}$ be a linear functional. Define the component $T^{\mu}_{\ \nu}$:

$$T^{\mu}_{\ \nu} \equiv T\left(\mathbf{e}^{(\mu)}, \boldsymbol{f}_{(\nu)}\right). \tag{8}$$

Let $\boldsymbol{\xi} \overset{(4)}{=} \xi_{\mu}\,\mathbf{e}^{(\mu)} \in \mathcal{V}^*$ and $\boldsymbol{W} \overset{(4)}{=} W^{\nu}\,\boldsymbol{f}_{(\nu)} \in \mathcal{W}^*$. Then

$$T(\boldsymbol{\xi}, \boldsymbol{W}) = T\left(\xi_{\mu}\,\mathbf{e}^{(\mu)},\ W^{\nu}\,\boldsymbol{f}_{(\nu)}\right) \overset{\text{(bilinear)}}{=} \xi_{\mu} W^{\nu}\, T\left(\mathbf{e}^{(\mu)},\ \boldsymbol{f}_{(\nu)}\right) \overset{(8)}{=} T^{\mu}_{\ \nu}\,\xi_{\mu} W^{\nu}$$

$$\overset{(6)}{=} T^{\mu}_{\ \nu}\, e^{\nu}_{\ \mu}(\boldsymbol{\xi}, \boldsymbol{W}) \quad \text{for all } (\boldsymbol{\xi}, \boldsymbol{W})$$

$$\Rightarrow \quad \boxed{T = T^{\mu}_{\ \nu}\, e^{\nu}_{\ \mu}} \quad ✔ \tag{9}$$

This concludes Step 2, to express a tensor in terms of a tensor basis, except that we have to finish proving that $\mathcal{B}_{\mathcal{T}}$ is indeed a basis.

The set $\left\{ e^{\nu}_{\ \mu} \right\}$ is linearly independent:

$$x^{\sigma}_{\ \rho}\, e^{\rho}_{\ \sigma} = 0 \quad \Rightarrow \quad x^{\sigma}_{\ \rho} = x^{\mu}_{\ \nu}\, \delta^{\sigma}_{\mu}\, \delta^{\nu}_{\rho} \overset{(4)}{=} x^{\mu}_{\ \nu}\, e^{\nu}_{\ \mu}\left(\mathbf{e}^{(\sigma)}, \boldsymbol{f}_{(\rho)}\right) = 0 \quad \forall\, \sigma, \rho \quad ✔$$

We have proven in passing that dim($\mathcal{T}$) = nm since that is the size of its basis.

Because of equation (9), we often loosely call $T^{\mu}_{\ \nu}$ a (1, 1) tensor.

<u>Step 3</u>: Define the tensor product '$\otimes$' of a vector V and a covector ω

Because vectors and covectors can be considered duals of each other, we can regard V as a linear functional:

$$V : \mathcal{V}^* \to \mathbf{F} : V(\xi) = \xi(\boldsymbol{V}) \overset{(4)}{=} \xi_{\nu} e^{(\nu)}\left(V^{\mu} \mathbf{e}_{(\mu)}\right) \overset{\text{(bilinear)}}{=} \xi_{\nu} V^{\mu} e^{(\nu)}\left(\mathbf{e}_{(\mu)}\right) \overset{(3)}{=} \xi_{\nu} V^{\mu} \delta^{\mu}_{\nu} = \xi_{\mu} V^{\mu} \in \mathbf{F}.$$

Of course, ω *is* a linear functional: $\omega : \mathcal{W} \to \mathbf{F} : \omega(\boldsymbol{W}) = \omega_{\nu} W^{\nu} \in \mathbf{F}$.

Definition The coordinate-independent definition of the tensor product is

$$V \otimes \omega : \mathcal{V}^* \times \mathcal{W} \to \mathbf{F} : \boxed{V \otimes \omega(\boldsymbol{\xi}, \boldsymbol{W}) = V(\xi)\,\omega(\boldsymbol{W})}. \tag{10}$$

Introducing coordinates, we can plug

$$V = e_{(\mu)}, \quad \omega = \varphi^{(\nu)}, \quad \boldsymbol{\xi} = \mathbf{e}^{(\sigma)}, \quad \text{and} \quad \boldsymbol{W} = \boldsymbol{f}_{(\rho)}$$

into (10) to get that for all $\mathbf{e}^{(\sigma)}$ and $\boldsymbol{f}_{(\rho)}$,

$$e_{(\mu)} \otimes \varphi^{(\nu)}\left(\mathbf{e}^{(\sigma)}, \boldsymbol{f}_{(\rho)}\right) \overset{(10)}{=} e_{(\mu)}\left(\mathbf{e}^{(\sigma)}\right) \varphi^{(\nu)}\left(\boldsymbol{f}_{(\rho)}\right) \overset{(3)}{=} \delta^{\sigma}_{\mu} \, \delta^{\nu}_{\rho} \overset{(5)}{=} e^{\nu}_{\mu}\left(\mathbf{e}^{(\sigma)}, \boldsymbol{f}_{(\rho)}\right).$$

That is.

$$\boxed{e_{(\mu)} \otimes \varphi^{(\nu)} = e^{\nu}_{\mu}}. \tag{11}$$

Claim: $\boxed{V \otimes \omega = V^{\mu} \, \omega_{\nu} \; e^{\nu}_{\mu}}$ (12)

$$V \otimes \omega = V^{\mu} e_{(\mu)} \otimes \omega_{\nu} \varphi^{(\nu)} \overset{\text{(bilinearity)}}{=} V^{\mu} \, \omega_{\nu} e_{(\mu)} \otimes \varphi^{(\nu)} \overset{(11)}{=} V^{\mu} \, \omega_{\nu} e^{\nu}_{\mu}.$$

Equation (11) shows that the tensor basis can be expressed as a tensor product of vector and covector bases, and equation (12) accomplishes our objective for Step 3. But, there is more to be learned. If we set $T^{\mu}_{\nu} = V^{\mu} \omega_{\nu}$ then $V \otimes \omega \overset{(12)}{=} T^{\mu}_{\nu} \; e^{\nu}_{\mu} \in \mathcal{T}$. That is, the tensor product in equation (12) is a rank (1,1) tensor. This inspires the following definition for the tensor product of a vector space with a covector space.

Definition $\boxed{\mathcal{V} \otimes \mathcal{W}^{*} \equiv \mathcal{T}}$.

The question arises as to whether the converse is true, that every rank (1,1) tensor can be expressed as a tensor product of a vector and covector. The definition of $\mathcal{V} \otimes \mathcal{W}^{*}$ suggests that it is true, and it is tempting to think so because the μ and ν on LHS of (11) match 1-1 with μ and ν on RHS. However, it turns out that tensor products $V \otimes \omega$ comprise only small part of the total set of rank (1, 1) tensors *T*. We give an example of a tensor *T* that is not equal to $V \otimes \omega$ for any vector *V* and covector ω.

Example: Let $\mathcal{V}$ and $\mathcal{W}$ have dimension *n* = 2 and define $T \overset{(9)}{=} T^{\mu}_{\nu} \, e^{\nu}_{\mu}$ by $T^{1}_{1} = T^{2}_{2} = 0$ and $T^{2}_{1} = \; T^{1}_{2} = 1$. Suppose $T = V \otimes \omega$. Since $T = T^{\mu}_{\nu} \, e^{\nu}_{\mu}$ and $V \otimes \omega \overset{(12)}{=} V^{\mu} \, \omega_{\nu} \, e^{\nu}_{\mu}$, then $T^{\mu}_{\nu} = V^{\mu} \omega_{\nu}$ for $\mu, \; \nu = 1, 2$. This represents 4 equations in the 4 unknowns V^1, V^2, ω_1 , and ω_2:

(1) $V^{1} \omega_{1} = \; 0$

(2) $V^{1} \omega_{2} = \; 1$

(3) $V^{2} \omega_{1} = \; 1$

(4) $V^{2} \omega_{2} = 0$

From (1), either $V^1 = 0$ or $\omega_1 = 0$. If $V^1 = 0$ then (2) yields 0 = 1. If $\omega_1 = 0$ then (3) yields 0 = 1. So, $T \neq V \otimes \omega$. ■

We conclude that not all rank (1,1) tensors are tensor products of vectors and covectors. This is somewhat odd since the rank (1, 1) tensor basis is composed exclusively of product tensors.

If $T^\mu_\nu \neq V^\mu \omega_\nu$ for any choice of V^μ and ω_ν, then, also, $T^\mu_\nu \neq T^\mu T_\nu$ for any choice of T^μ and T_ν. That is, T^μ_ν cannot in general be decomposed into $T^\mu T_\nu$.

In Appendix A we give an example very much like the example above to show that if you add two product tensors, the result might not be another product tensor. It will, of course, be a (1, 1) tensor. The (1, 1) tensors T in $\mathcal{T} = \mathcal{V} \otimes \mathcal{W}^*$ are sums of product tensors $V \otimes \omega$.

Step 4: Show that a rank (1,1) tensor can be represented by a matrix.

We start with the subset of tensors that are tensor products. In Example 2 of Section 3 we saw that we can represent the tensor product $V \otimes \omega$ as a matrix:

$$V \otimes \omega \leftrightarrow \begin{pmatrix} V^1 \\ \vdots \\ V^n \end{pmatrix} \otimes \left(\omega_1 \cdots \omega_m\right) = \begin{pmatrix} & \vdots & \\ \cdots & V^\mu \otimes \omega_\nu & \cdots \\ & \vdots & \end{pmatrix}. \tag{13}$$

More generally, a rank (1, 1) tensor $T = T^\mu_\nu e^\nu_\mu$ can be represented as a matrix by

$$T \leftrightarrow T^\mu_\nu \boldsymbol{E}^\nu_\mu \overset{(3.3)}{=} \begin{pmatrix} T^1_1 & \cdots & T^1_m \\ \vdots & & \vdots \\ T^n_1 & \cdots & T^n_m \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & T^\mu_\nu & \cdots \\ & \vdots & \end{pmatrix}. \tag{14}$$

This shows that $\mathcal{T}$ is isomorphic to the vector space of nxn matrices.

Step 5: Show how to express $T(\xi, \boldsymbol{W})$ in terms of matrix algebra

$$\boxed{T(\xi, \boldsymbol{W}) = T^\mu_\nu \xi_\mu W^\nu}: \tag{15}$$

$$T(\xi, \boldsymbol{W}) \overset{(4)}{=} T\left(\xi_\mu \mathbf{e}^{(\mu)}, W^\nu \boldsymbol{f}_{(\nu)}\right) \overset{\text{(bilinearity)}}{=} \xi_\mu W^\nu T\left(\mathbf{e}^{(\mu)}, \boldsymbol{f}_{(\nu)}\right) \overset{(8)}{=} T^\mu_\nu \xi_\mu W^\nu$$

Therefore,

$$T\left(\boldsymbol{\xi},\boldsymbol{W}\right)\overset{(15)}{=}T^{\mu}_{\nu}\,\xi_{\mu}\,W^{\nu}=\left(\xi_1\cdots\xi_n\right)\begin{pmatrix} T^1_1 & \cdots & T^1_m \\ \vdots & & \vdots \\ T^n_1 & \cdots & T^n_m \end{pmatrix}\begin{pmatrix} W^1 \\ \vdots \\ W^m \end{pmatrix} \tag{16}$$

as is easily confirmed because when the the matrix multiplication is carried out, the result equals RHS of equation (15).

Step 6: Show that $T\left(\boldsymbol{\xi},\boldsymbol{W}\right)$ can operate as a function of only one variable

Suppose $T\left(\boldsymbol{\xi},\boldsymbol{W}\right)=T^{\mu}_{\nu}\,\xi_{\mu}\,W^{\nu}$. We can think of T as a linear functional on $\mathcal{W}$ as follows. Define $T\left(\boldsymbol{W}\right)\equiv T^{\mu}_{\nu}\,W^{\nu}$. Then $T:\mathcal{W}\to\mathbf{F}$, and it is straightforward to show that T is linear. So, although T^{μ}_{ν} is a rank (1, 1) tensor, we can think of $T^{\mu}_{\nu}\,W^{\nu}$ as a rank (0, 1) tensor. Similarly, if $T\left(\boldsymbol{\xi}\right)\equiv T^{\mu}_{\nu}\,\xi_{\mu}$, then T is a linear functional on $\mathcal{V}^*$ and we can think of $T^{\mu}_{\nu}\,\xi_{\mu}$ as a rank (1, 0) tensor.

Step 7: Define the tensor product '$\otimes$' of two rank (1,1) tensors

Not only can we take the tensor product of vectors and covectors, we can take the tensor product of two tensors, even if the tensors have different domains. For example, suppose $R:\mathcal{V}^*\times\mathcal{W}\to\mathbf{F}$: $R\overset{(9)}{=}R^{\mu}_{\nu}\,e^{\nu}_{\mu}$ and $S:\mathcal{X}^*\times\Upsilon\to\mathbf{F}$: $S\overset{(9)}{=}S^{\sigma}_{\rho}\,f^{\rho}_{\sigma}$.
Define a (2, 2) tensor

$$R\otimes S:\mathcal{V}^*\times\mathcal{X}^*\times\mathcal{W}\times\Upsilon\to\mathbf{F}:\ R\otimes S\left(\boldsymbol{\xi},\boldsymbol{\chi},\boldsymbol{W},\boldsymbol{Y}\right)=R\left(\boldsymbol{\xi},\boldsymbol{W}\right)S\left(\boldsymbol{\chi},\boldsymbol{Y}\right). \tag{17}$$

Note that we have swapped $\mathcal{X}^*$ and $\mathcal{W}$ from the domains of R and S in order to make the domain of $R\otimes S$ match the requirement of equation (1) that covector spaces precede vector spaces. Also note that this is the coordinate-free, preferred type of definition.

If we set $T\equiv R\otimes S$, then we can write the coordinate-dependent equation

$$T=T^{\mu\sigma}_{\nu\rho}\,e^{\nu\rho}_{\mu\sigma} \tag{18}$$

where

$$T^{\mu\sigma}_{\nu\rho}\equiv R^{\mu}_{\nu}S^{\sigma}_{\rho}\ \text{ and }\ e^{\nu}_{\mu}\otimes f^{\rho}_{\sigma}=e^{\nu\rho}_{\mu\sigma}. \tag{19}$$

We leave the details of defining the linear functional $e^{\nu\rho}_{\mu\sigma}$ to Section 6.

Armed with a basis, so that

$$R\left(\boldsymbol{\xi},\boldsymbol{W}\right)\overset{(15)}{=}R^{\mu}_{\nu}\,\xi_{\mu}W^{\nu} \text{ and } S\left(\boldsymbol{\chi},\boldsymbol{Y}\right)\overset{(15)}{=}S^{\sigma}_{\rho}\,\chi_{\sigma}Y^{\rho} \ ,$$

we can write

$$R\otimes S\left(\boldsymbol{\xi},\boldsymbol{\chi},\boldsymbol{W},\boldsymbol{Y}\right)\overset{(17)}{=}R\left(\boldsymbol{\xi},\boldsymbol{W}\right)S\left(\boldsymbol{\chi},\boldsymbol{Y}\right)=R^{\mu}_{\nu}\,S^{\sigma}_{\rho}\xi_{\mu}\,\chi_{\sigma}\,W^{\nu}Y^{\rho} \ . \tag{20}$$

Finally, the standalone formula for $R\otimes S$ can be expressed as

$$R\otimes S=R^{\mu}_{\nu}S^{\sigma}_{\ \rho}e^{\mu\sigma}_{\nu\rho}\,. \tag{21}$$

This equation is a special case of equation (19) in Section 6, and its proof is deferred until then.

6 Rank (k, ℓ) Tensors

We can develop the equations of rank (k, ℓ) tensors with each vector space $\mathcal{V}_i$ and covector space $\mathcal{W}_j^*$ distinct as we did in Section 5, but the notation quickly explodes to unreadability. Therefore, we now assume that all of the vector and dual spaces are copies from a single vector space $\mathcal{V}$ of dimension n; i.e.,

$$\begin{aligned} &\mathcal{V}_i = \mathcal{V} \ \text{ for } i = 1,\cdots,k \qquad &&\mathcal{W}_j = \mathcal{V} \ \text{ for } j = 1,\cdots,\ell \\ &\mathcal{V}_i^* = \mathcal{V}^* \ \text{ for } i = 1,\cdots,k \qquad &&\mathcal{W}_j^* = \mathcal{V}^* \ \text{ for } j = 1,\cdots,\ell \end{aligned}.$$

We develop rank (k, ℓ) tensors mirroring the steps used in Section 5 for rank (1, 1) tensors. Proofs are omitted since they were provided in Section 5. We use the index convention

$i = 1,\cdots,k$ for the k vector spaces $\mathcal{V}_i$ and the k covector spaces $\mathcal{V}_i^*$

$j = 1,\cdots,\ell$ for the ℓ vector spaces $\mathcal{W}_j$ and the ℓ covector spaces $\mathcal{W}_j^*$

$\mu, \nu, \sigma, \rho = 1,\cdots,n$ for the indices on components of vectors and covectors

Moreover, any of these Greek indices might have subscripts. For example, a vector $\boldsymbol{V}_{(i)}$ in vector space $\mathcal{V}_i$ will have components denoted

$$V_{(i)}^{\mu_i}, \quad \mu_i = 1,\cdots,n.$$

<u>Step 1</u>: Define rank (k, ℓ) tensors

$$\mathcal{T} \equiv \left\{ T : \mathcal{V}_1^* \times \cdots \times \mathcal{V}_k^* \times \mathcal{W}_1 \times \cdots \times \mathcal{W}_\ell \to \mathbf{F} : \ T \text{ is multilinear} \right\}. \tag{1}$$

To simplify expression (1), set

$$\mathcal{V}^* = \mathcal{V}_1^* \times \cdots \times \mathcal{V}_k^* \ \text{ and } \quad \mathcal{W} = \mathcal{W}_1 \times \cdots \times \mathcal{W}_\ell, \tag{1A}$$

This enables us to rewrite (1) to more closely resemble how we expressed (1) for (1, 1) tensors in Section 5:

$$\mathcal{T} \equiv \left\{ T : \mathcal{V}^* \times \mathcal{W} \to \mathbf{F} : \ T \text{ is multilinear} \right\}. \tag{1'}$$

We also introduce the notation

$$\mathcal{V}^k \equiv \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_k \qquad \mathcal{W}_\ell^* \equiv \mathcal{W}_1^* \otimes \cdots \otimes \mathcal{W}_\ell^* \qquad \mathcal{T} \equiv \mathcal{V}_k^\ell \equiv \mathcal{V}^k \otimes \mathcal{W}_\ell^* \tag{1B}$$

Members of $\mathcal{V}^k$ are called **contravariant rank (k, 0) tensors**. Members of $\mathcal{W}_\ell^*$ are called **rank (0, ℓ) covariant tensors**. Members of $\mathcal{V}_k^\ell$ are called **mixed rank (k, ℓ) tensors**. We can rewrite (1) again as

$$\mathcal{V}^k \otimes \mathcal{W}_\ell{}^* \equiv \left\{ T : \mathcal{V}^* \times \mathcal{W} \to \mathbf{F} : \; T \text{ is multilinear} \right\} \tag{1''}$$

Observe the reversal of the * symbol because vectors operate on covectors and vice-versa. Importantly, we cannot replace *T* by $V \otimes \omega$ because *T* is a sum of product tensors and cannot in general be represented by a single product tensor.

<u>Step 2</u>: Express tensor *T* in terms of a tensor basis

Since there is only one reference vector space $\mathcal{V}$, we specify just a pair of bases:

$$\text{Bases:} \quad \mathcal{B} = \left\{ e_{(\mu)} \right\}_{\mu=1}^{n} \text{ and } \mathcal{B}^* = \left\{ e^{(\nu)} \right\}_{\nu=1}^{n} \tag{2}$$

where the bases satisfy the covector relationships, equations (1.1) and (1.2):

$$e^{(\rho)}\left(\mathbf{e}_{(\mu)}\right) = \delta^{\rho}_{\mu} \text{ and } e_{(\mu)}\left(\mathbf{e}^{(\sigma)}\right) = \delta^{\sigma}_{\mu}. \tag{3}$$

The basis for $\mathcal{V}^* \times \mathcal{W}$ is then expressed:

$$\mathcal{B}_{\mathcal{V}^* \times \mathcal{W}} = \left\{ e^{(\nu_1)}, \cdots, e^{(\nu_k)}, e_{(\rho_1)}, \cdots, e_{(\rho_\ell)} \right\} \; .$$

Next we denote a pair of vectors in $\mathcal{V}$ and a pair of covectors in $\mathcal{V}^*$:

Vectors: $\quad V = V^{\mu} e_{(\mu)} \in \mathcal{V}, \;\; \boldsymbol{W} = W^{\rho} \mathbf{e}_{(\rho)} \in \mathcal{V},$

$$\text{Covectors} \quad \boldsymbol{\xi} = \xi_{\nu} e^{(\nu)} \in \mathcal{V}^*, \text{ and } \omega = \omega_{\sigma} \mathbf{e}^{(\sigma)} \in \mathcal{V}^* \tag{4}$$

More specifically,

A vector in $\mathcal{V}_i$ will be expressed as $\quad V_{(i)} = V^{\mu_i}_{(i)} e_{(\mu_i)}$

A vector in $\mathcal{W}_j$ will be expressed as $\quad \boldsymbol{W}_{(j)} = W^{\rho_j}_{(j)} \mathbf{e}_{(\rho_j)}$

A covector in $\mathcal{V}_i{}^*$ will be expressed as $\quad \boldsymbol{\xi}^{(i)} = \xi^{(i)}_{\nu_i} \mathbf{e}^{(\nu_i)}$

$$\text{A covector in } \mathcal{W}_j{}^* \text{ will be expressed as} \quad \omega^{(j)} = \omega^{(j)}_{\sigma_j} e^{(\sigma_j)} \tag{4'}$$

Now, define $e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}$, the basis for $\mathcal{T}'$, first by defining it as a function on $\mathcal{B}_{\mathcal{V}^* \times \mathcal{W}}$ and then extending it multilinearly to $\mathcal{V}^* \times \mathcal{W}$:

$$\boldsymbol{e}^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}\left(\mathbf{e}^{(\sigma_1)}, \cdots, \mathbf{e}^{(\sigma_k)}, \mathbf{e}_{(\rho_1)}, \cdots, \mathbf{e}_{(\rho_\ell)}\right) \equiv \delta^{\sigma_1}_{\mu_1} \cdots \delta^{\sigma_k}_{\mu_k} \, \delta^{\nu_1}_{\rho_1} \cdots \delta^{\nu_\ell}_{\rho_\ell}. \tag{5}$$

$$e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k} : \mathcal{V}^* \times \mathcal{W} \to \mathbf{F} : e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}\left(\boldsymbol{\xi}^{(1)}, \cdots, \boldsymbol{\xi}^{(k)}, \boldsymbol{W}_{(1)}, \cdots, \boldsymbol{W}_{(\ell)} \right) = \xi^{(1)}_{\mu_1} \cdots \xi^{(k)}_{\mu_k} W^{\nu_1}_{(1)} \cdots W^{\nu_\ell}_{(\ell)} \tag{6}$$

$$\mathcal{B}_{\mathcal{T}} = \left\{ e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k} : \mu_i = 1, \cdots, n\, ;\ \nu_j = 1, \cdots, n \right\} \tag{7}$$

Express T as a linear combination of basis vectors:

$$\boxed{T = T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\ e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}} \tag{8}$$

where

$$T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} = T\left(\boldsymbol{e}^{(\mu_1)}, \cdots, \boldsymbol{e}^{(\mu_k)}, \boldsymbol{e}_{(\nu_1)}, \cdots, \boldsymbol{e}_{(\nu_\ell)} \right). \tag{9}$$

From (7), the basis has size $n^{k+\ell}$. Hence we have discovered that

$\dim(\mathcal{V}^k \otimes \mathcal{W}_\ell{}^*) = n^{k+\ell}$.

Because of equation (8), we often loosely refer to $T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}$ as a (k, ℓ) tensor.

Step 3: Define the tensor product '$\otimes$' of vectors and covectors

Definition The coordinate-independent definition of the tensor product is

$$V_{(1)} \otimes \cdots \otimes V_{(k)} \otimes \omega^{(1)} \otimes \cdots \otimes \omega^{(\ell)} : \mathcal{V}_1{}^* \times \cdots \times \mathcal{V}_k{}^* \times \mathcal{W}_1 \times \cdots \times \mathcal{W}_\ell \to \mathbf{F}:$$

$$\boxed{\begin{aligned} &V_{(1)} \otimes \cdots \otimes V_{(k)} \otimes \omega^{(1)} \otimes \cdots \otimes \omega^{(\ell)} \left(\boldsymbol{\xi}^{(1)}, \cdots, \boldsymbol{\xi}^{(k)}, \boldsymbol{W}_{(1)}, \cdots, \boldsymbol{W}_{(\ell)} \right) \\ &= V_{(1)}\left(\boldsymbol{\xi}^{(1)} \right) \cdots V_{(k)}\left(\boldsymbol{\xi}^{(k)} \right)\ \omega^{(1)}\left(\boldsymbol{W}_{(1)} \right) \cdots \omega^{(\ell)}\left(\boldsymbol{W}_{(\ell)} \right) \end{aligned}}. \tag{10}$$

where, by regarding $\boldsymbol{V}_{(i)}$ as a dual vector of $\mathcal{V}_{(i)}{}^*$, we have $V_{(i)}\left(\boldsymbol{\xi}^{(i)} \right) = \xi^{(i)}\left(\boldsymbol{V}_{(i)} \right)$.

To generate the coordinate-dependent equation, first plug into (10) the values

$$V_{(i)} = e_{(\mu_i)},\quad \omega^{(j)} = e^{(\nu_j)},\quad \boldsymbol{\xi}^{(i)} = \boldsymbol{e}^{(\sigma_i)},\ \text{ and }\ \boldsymbol{W}_{(j)} = \boldsymbol{e}_{(\rho_j)}\ ,$$

as in Step 3 of Section 5, to get

$$\boxed{e_{(\mu_1)} \otimes \cdots \otimes e_{(\mu_k)} \otimes e^{(\nu_1)} \otimes \cdots \otimes e^{(\nu_\ell)} = e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}} \tag{11}$$

and then get

$$\boxed{V_{(1)} \otimes \cdots \otimes V_{(k)} \otimes \omega^{(1)} \otimes \cdots \otimes \omega^{(\ell)} \equiv V^{\mu_1}_{(1)} \cdots V^{\mu_k}_{(k)}\, \omega^{(1)}_{\nu_1} \cdots \omega^{(\ell)}_{\nu_\ell}\ e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}}. \tag{12}$$

Equation (11) shows that the tensor basis can be expressed as a tensor product of vector and covector bases.

If we set $T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell} = V^{\mu_1}_{(1)}\cdots V^{\mu_k}_{(k)}\omega^{(1)}_{\nu_1}\cdots\omega^{(\ell)}_{\nu_\ell}$ in (12), then

$$V_{(1)}\otimes\cdots\otimes V_{(k)}\otimes\omega^{(1)}\otimes\cdots\otimes\omega^{(\ell)} = T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}e^{\nu_1\cdots\nu_\ell}_{\mu_1\cdots\mu_k}\in\mathcal{T}$$

That is, the tensor product of k vectors and ℓ covectors is a (k,ℓ) tensor.

However, as shown in Step 3 inSection 5, the converse is false.
Not all rank (k,ℓ) tensors are tensor products. They are, in general, sums of tensor products.

Technical Note Equation (12) represents $V\otimes W\otimes\omega$, $V\otimes W$, and $W\otimes\omega$. To be consistent, we must have that $(V\otimes W)\otimes\omega = (V\otimes W)\otimes\omega = V\otimes(W\otimes\omega)$.

$$(V\otimes W)\otimes\omega \overset{(12)}{=} V^\mu W^\rho e_{\mu\rho}\otimes\omega_\nu e^{(\nu)} \overset{\text{(multilinear)}}{=} V^\mu W^\rho\omega_\nu\, e_{\mu\rho}\otimes e^{(\nu)}$$

and

$$V\otimes(W\otimes\omega) \overset{(12)}{=} V^\mu W^\rho\omega_\nu e_{(\mu)}\otimes e^\rho_\nu\ .$$

However, $e_{\mu\rho}\otimes e^{(\nu)}$ and $e_{(\mu)}\otimes e^\rho_\nu$, on the right-hand sides above, have not yet been defined. We define them now.

Definition $e^{\nu_1\cdots\nu_\ell}_{\mu_1\cdots\mu_k}\otimes e^{\sigma_1\cdots\sigma_n}_{\rho_1\cdots\rho_m}\equiv e^{\nu_1\cdots\nu_\ell\,\sigma_1\cdots\sigma_n}_{\mu_1\cdots\mu_k\,\rho_1\cdots\rho_m}$ (13)

Using this definition,

$$(V\otimes W)\otimes\omega = V^\mu W^\rho\omega_\nu e_{\mu\rho}\otimes e^\nu \overset{(13)}{=} V^\mu W^\rho\omega_\nu e^\nu_{\mu\rho} \overset{(12)}{=} V\otimes W\otimes\omega$$

and

$$V\otimes(W\otimes\omega) = V^\mu W^\rho\omega_\nu \mathbf{e}_{(\mu)}\otimes e^\nu_\rho \overset{(13)}{=} V^\mu W^\rho\omega_\nu e^\nu_{\mu\rho} \overset{(12)}{=} V\otimes W\otimes\omega \qquad ✔$$

Step 4: Show that a rank (k, ℓ) tensor can be represented by a matrix

We begin with the subset of tensors that are tensor products. We use the pattern definition introduced in Section 3. The product of k vectors and ℓ covectors can be represented by an $n^k\times n^\ell$ matrix:

$$V_{(1)} \otimes \cdots \otimes V_{(k)} \otimes \omega^{(1)} \otimes \cdots \otimes \omega^{(\ell)}$$

$$\leftrightarrow \begin{pmatrix} V_{(1)}^{1} \\ \vdots \\ V_{(1)}^{\mu_1} \\ \vdots \\ V_{(1)}^{n} \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} V_{(k)}^{1} \\ \vdots \\ V_{(k)}^{\mu_k} \\ \vdots \\ V_{(k)}^{n} \end{pmatrix} \otimes \left(\omega_1^{(1)} \cdots \omega_{\nu_1}^{(1)} \cdots \omega_n^{(1)} \right) \otimes \cdots \otimes \left(\omega_1^{(\ell)} \cdots \otimes \omega_{\nu_\ell}^{(\ell)} \otimes \omega_n^{(\ell)} \right) \tag{14}$$

$$= \begin{pmatrix} & \vdots & \\ \cdots & V_{(1)}^{\mu_1} \otimes \cdots \otimes V_{(k)}^{\mu_k} \otimes \omega_{\nu_1}^{(1)} \otimes \cdots \otimes \omega_{\nu_\ell}^{(\ell)} & \cdots \\ & \vdots & \end{pmatrix}$$

Similarly, a more general rank **(**k, ℓ) tensor can be represented by the $n^k \times n^\ell$ matrix

$$T \overset{(8)}{=} T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k} \leftrightarrow \begin{pmatrix} & \vdots & \\ \cdots & T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} & \cdots \\ & \vdots & \end{pmatrix} \equiv \boldsymbol{M}_T \,. \tag{15}$$

This shows that $\mathcal{V}^k \otimes \mathcal{W}_\ell{}^*$ is isomorphic to the vector space of $n^k \times n^\ell$ matrices.

<u>Step 5</u>: Show how to express $T\left(\boldsymbol{\xi}^{(1)},\cdots,\boldsymbol{\xi}^{(k)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)}\right)$ in terms of matrix algebra

Mirroring Step 5 in Section 5, we develop the abstract equation

$$\boxed{T\left(\boldsymbol{\xi}^{(1)},\cdots,\boldsymbol{\xi}^{(k)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)}\right) = T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\ \xi^{(1)}_{\mu_1} \cdots \xi^{(k)}_{\mu_k} W^{\nu_1}_{(1)} \cdots W^{\nu_\ell}_{(\ell)}} \,. \tag{16}$$

RHS of (16) can be represented by a matrix as follows. Define

row vector: $\boldsymbol{r}_\xi = \left(\xi_1^{(1)} \cdots \xi_1^{(k)} \quad \cdots \quad \xi_n^{(1)} \cdots \xi_n^{(k)} \right)$,

and

column vector: $\boldsymbol{C_W} = \begin{pmatrix} W^1_{(1)} \cdots W^1_{(\ell)} \\ \vdots \\ W^n_{(1)} \cdots W^n_{(\ell)} \end{pmatrix}.$

Also, matrix $\boldsymbol{M}_T$ was defined in equation (15).

Claim

$$T\left(\boldsymbol{\xi}^{(1)},\cdots,\boldsymbol{\xi}^{(k)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)}\right) = \boldsymbol{r}_\xi\,\boldsymbol{M}_T\,\boldsymbol{C_W}$$

$$= \left(\xi^{(1)}_1 \cdots \xi^{(k)}_1 \quad \cdots \quad \xi^{(1)}_n \cdots \xi^{(k)}_n\right) \begin{pmatrix} T^{1\cdots1}_{1\cdots1} & \cdots & T^{1\cdots1}_{n\cdots n} \\ \vdots & & \vdots \\ T^{n\cdots n}_{1\cdots1} & \cdots & T^{n\cdots n}_{n\cdots n} \end{pmatrix} \begin{pmatrix} W^1_{(1)} \cdots W^1_{(\ell)} \\ \vdots \\ W^n_{(1)} \cdots W^n_{(\ell)} \end{pmatrix}: \tag{17}$$

We prove Equation (17) by multiplying it out. It is easier to perform the multiplication if we insert a middle term into each of the three matrix objects. We multiply the row times the matrix first, and we focus on the middle terms.

$$\left(\xi^{(1)}_1 \cdots \xi^{(k)}_1 \ \cdots \ \xi^{(1)}_{\mu_1} \cdots \xi^{(i)}_{\mu_i} \cdots \xi^{(k)}_{\mu_k} \ \cdots \ \xi^{(1)}_n \cdots \xi^{(k)}_n\right) \begin{pmatrix} T^{1\cdots1}_{1\cdots1} & \cdots & T^{1\cdots1}_{\nu_1\cdots\nu_\ell} & \cdots & T^{1\cdots1}_{n\cdots n} \\ \vdots & & \vdots & & \vdots \\ T^{\mu_1\cdots\mu_k}_{1\cdots1} & \cdots & T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell} & \cdots & T^{\mu_1\cdots\mu_k}_{n\cdots n} \\ \vdots & & \vdots & & \vdots \\ T^{n\cdots n}_{1\cdots1} & \cdots & T^{n\cdots n}_{\nu_1\cdots\nu_\ell} & \cdots & T^{n\cdots n}_{n\cdots n} \end{pmatrix} \boldsymbol{C_W}$$

$$= \left(\cdots \qquad T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}\xi^{(1)}_{\mu_1}\cdots\xi^{(i)}_{\mu_i}\cdots\xi^{(k)}_{\mu_k} \qquad \cdots\right) \begin{pmatrix} W^1_{(1)} \cdots W^1_{(\ell)} \\ \vdots \\ W^{\nu_1}_{(1)} \cdots W^{\nu_j}_{(j)} \cdots W^{\nu_\ell}_{(\ell)} \\ \vdots \\ W^n_{(1)} \cdots W^n_{(\ell)} \end{pmatrix}$$

$$= T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}\xi^{(1)}_{\mu_1}\cdots\xi^{(i)}_{\mu_i}\cdots\xi^{(k)}_{\mu_k} W^{\nu_1}_{(1)}\cdots W^{\nu_j}_{(j)}\cdots W^{\nu_\ell}_{(\ell)} = \text{RHS}(16) \qquad ✔$$

<u>Step 6</u>: Show that $T\left(\xi^{(1)},\cdots,\xi^{(k)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)}\right)$ can operate as a function of any subset of its arguments

We illustrate a couple of examples. $T\left(\xi^{(1)},\boldsymbol{W}_{(1)}\right)$ would be computed as

$$T\left(\xi^{(1)},\boldsymbol{W}_{(1)}\right)\equiv T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}\ \xi^{(1)}_{\mu_1}W^{\nu_1}_{(1)}\in\mathbf{F},$$

where

$$T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}\overset{(9)}{=}T\left(\boldsymbol{e}^{(\mu_1)},\cdots,\boldsymbol{e}^{(\mu_k)},\boldsymbol{e}_{(\nu_1)},\cdots,\boldsymbol{e}_{(\nu_\ell)}\right).$$

This shows that T can operate as a (1, 1) tensor. Similarly, if we denote

$$T\left(\xi^{(2)},\cdots,\xi^{(k)},\boldsymbol{W}_{(2)},\cdots,\boldsymbol{W}_{(\ell)}\right)\equiv T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}\ \xi^{(2)}_{\mu_2}\cdots\xi^{(k)}_{\mu_k}W^{\nu_2}_{(2)}\cdots W^{\nu_\ell}_{(\ell)}$$

then T operates as a (k-1, ℓ-1) tensor.

More often we work only with the components. In that case, loosely speaking, $S^{\mu_2\cdots\mu_k}_{\nu_2\cdots\nu_\ell}\equiv T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}\ \xi^{(1)}_{\mu_1}W^{\nu_1}_{(1)}$ is a (k-1, ℓ-1) tensor and $R^{\mu_1}_{\nu_1}\equiv T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_\ell}\ \xi^{(2)}_{\mu_2}\cdots\xi^{(k)}_{\mu_k}W^{\nu_2}_{(2)}\cdots W^{\nu_\ell}_{(\ell)}$ is a (1, 1) tensor. More properly speaking, $S=S^{\mu_2\cdots\mu_k}_{\nu_2\cdots\nu_\ell}\ e^{\nu_2\cdots\nu_\ell}_{\mu_2\cdots\mu_k}$ is a (k-1, ℓ-1) tensor and $R=R^{\mu_1}_{\nu_1}e^{\nu_1}_{\mu_1}$ is a (1, 1) tensor.

<u>Step 7</u>: Define the tensor product '$\otimes$' of a rank (k, ℓ) tensor and a rank (r ,s) tensor

Not only can we form the tensor product of vectors and covectors, we can form the tensor product $R\otimes S$ of two tensors even if the tensors have different ranks.

Definition If $R:\mathcal{V}^*\times\mathcal{W}\to\mathbf{F}$ is a (k, ℓ) tensor and $S:\mathcal{X}^*\times\mathcal{Y}\to\mathbf{F}$ is an (r, s) tensor, then the coordinate-free definition of the (k+r, ℓ+s) tensor $R\otimes S$ is

$$R\otimes S:\mathcal{V}^*\times:\mathcal{X}^*\times\mathcal{W}\times\mathcal{Y}\to\mathbf{F}:$$

$$\boxed{\begin{aligned}&R\otimes S\left(\xi^{(1)},\cdots,\xi^{(k)},\chi^{(1)},\cdots,\chi^{(r)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)},\boldsymbol{Y}_{(1)},\cdots,\boldsymbol{Y}_{(s)}\right)\\&=R\left(\xi^{(1)},\cdots,\xi^{(k)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)}\right)S\left(\chi^{(1)},\cdots,\chi^{(r)},\boldsymbol{Y}_{(1)},\cdots,\boldsymbol{Y}_{s}\right)\end{aligned}}\tag{17}$$

Observe that the covector arguments of S have been grouped with the covector arguments of R in order for equation (17) to be compatible with definition (1) of a tensor.

When coordinates are introduced, we assume that

$$\mathcal{X}_{(1)} = \cdots = \mathcal{X}_{(r)} = \Upsilon_{(1)} = \cdots = \Upsilon_{(s)} = \mathcal{V}$$

and that

$\left\{ e_{\mu} \right\}_{\mu=1}^{n}$ is the basis for each of them.

Then, the coordinate-dependent formula for $R \otimes S$, using (17), is

$$\boxed{\begin{aligned} &R \otimes S\left(\boldsymbol{\xi}^{(1)}, \cdots, \boldsymbol{\xi}^{(k)}, \boldsymbol{\chi}^{(1)}, \cdots, \boldsymbol{\chi}^{(r)}, \boldsymbol{W}_{(1)}, \cdots, \boldsymbol{W}_{(\ell)}, \boldsymbol{Y}_{(1)}, \cdots, \boldsymbol{Y}_{(s)} \right) \\ &= R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\ S^{\rho_1 \cdots \rho_s}_{\sigma_1 \cdots \sigma_r}\ \xi^{(1)}_{\mu_1} \cdots \xi^{(k)}_{\mu_k} \chi^{(1)}_{\rho_1} \cdots \chi^{(r)}_{\rho_r} W^{\nu_1}_{(1)} \cdots W^{\nu_\ell}_{(\ell)} Y^{\sigma_1}_{(1)} \cdots Y^{\sigma_s}_{(s)} \end{aligned}}. \tag{18}$$

This is because

$$R\left(\boldsymbol{\xi}^{(1)}, \cdots, \boldsymbol{\xi}^{(k)}, \boldsymbol{W}_{(1)}, \cdots, \boldsymbol{W}_{(\ell)} \right) \overset{(16)}{=} R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\ \xi^{(1)}_{\mu_1} \cdots \xi^{(k)}_{\mu_k} W^{\nu_1}_{(1)} \cdots W^{\nu_\ell}_{(\ell)}$$

and

$$S\left(\chi^{(1)}, \cdots, \chi^{(r)}, \boldsymbol{Y}_{(1)}, \cdots, \boldsymbol{Y}_{(s)} \right) = S^{\rho_1 \cdots \rho_r}_{\sigma_1 \cdots \sigma_s}\ \chi^{(1)}_{\rho_1} \cdots \chi^{(r)}_{\rho_r} Y^{\sigma_1}_{(1)} \cdots Y^{\sigma_s}_{(s)}.$$

Lastly, to find the standalone formula for $R \otimes S$, we start with

$$R \overset{(8)}{=} R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\ e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k} \quad \text{and} \quad S = S^{\rho_1 \cdots \rho_r}_{\sigma_1 \cdots \sigma_s}\ e^{\sigma_1 \cdots \sigma_s}_{\rho_1 \cdots \rho_r}.$$

Then,

$$\boxed{R \otimes S = R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\ S^{\rho_1 \cdots \rho_r}_{\sigma_1 \cdots \sigma_s}\ e^{\nu_1 \cdots \nu_\ell \sigma_1 \cdots \sigma_s}_{\mu_1 \cdots \mu_k \rho_1 \cdots \rho_r}}. \tag{19}$$

This is due to the fact that

$$e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k} \otimes e^{\sigma_1 \cdots \sigma_s}_{\rho_1 \cdots \rho_r} \overset{(13)}{=} e^{\nu_1 \cdots \nu_\ell \sigma_1 \cdots \sigma_s}_{\mu_1 \cdots \mu_k \rho_1 \cdots \rho_r}.$$

Appendix A

The collection $\{V \otimes \omega\}$ of tensor products of vectors with covectors does not form a vector space. In fact, $V \otimes \xi \; + \; W \otimes \omega$ might not belong to the set.

Counter Example. Let indices μ and ν range from 1 to 2. Suppose there is a vector *x* and a covector ψ such that

$$x \otimes \psi = V \otimes \xi \; + \; W \otimes \omega = V^{\mu} \xi_{\nu} e^{\nu}_{\mu} + W^{\mu} \omega_{\nu} e^{\nu}_{\mu} = \left(V^{\mu} \xi_{\nu} + W^{\mu} \omega_{\nu}\right) e^{\nu}_{\mu}. \qquad (1)$$

Since

$$x = x^{\mu} e_{(\mu)} \text{ and } \psi = \psi_{\nu} e^{(\nu)},$$

then

$$x \otimes \psi = x^{\mu} \psi_{\nu} e^{\nu}_{\mu}.$$

So,

$$V^{\mu} \xi_{\nu} + W^{\mu} \omega_{\nu} = x^{\mu} \psi_{\nu} \text{ for } \mu = 1, 2 \text{ and } \nu = 1, 2.$$

This represents 4 equations in the four unknowns x^{μ} and ψ_{ν}.

$$\begin{cases} V^1 \xi_1 & + & W^1 \omega_1 & = & x^1 \psi_1 \\ V^1 \xi_2 & + & W^1 \omega_2 & = & x^1 \psi_2 \\ V^2 \xi_1 & + & W^2 \omega_1 & = & x^2 \psi_1 \\ V^2 \xi_2 & + & W^2 \omega_2 & = & x^2 \psi_2 \end{cases}.$$

Now suppose $V^1 = V^2 = W^1 = \xi_1 = \xi_2 = \omega_2 = 1$ and $\omega_1 = W^2 = -1$. Then we have

$$\begin{cases} 1 & - & 1 & = & 0 & = & x^1 \psi_1 \\ 1 & + & 1 & = & 2 & = & x^1 \psi_2 \\ 1 & + & 1 & = & 2 & = & x^2 \psi_1 \\ 1 & - & 1 & = & 0 & = & x^2 \psi_2 \end{cases} \Rightarrow \quad x^1 = 0 \text{ or } \xi_1 = 0.$$

If $x^1 = 0$ then equation 2 is false (i.e., 2 = 0). If $\xi_1 = 0$ then equation 3 is false.

Therefore there is no vector *x* and covector ψ that satisfy (1). So the collection is not closed under addition and thus is not a vector space. ■