Introduction to Tensors

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.

Tensors are defined in terms linear functionals. Tensors are manipulated as abstract symbols. Computations among tensors are often carried out using matrices.

For this reason, we begin by providing the requisite background in vectors spaces, linear functionals, and matrices.

1 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

The vector space definition also includes several rules that ensure proper behavior of addition and scalar multiplication, such as forming an Abelian group under addition, that will not be listed here.

Definition **Scalars** are simply elements of **F**. For quantum mechanics, the field is the complex numbers, $\mathbb{C}$. For spacetime, $\mathbb{R}$ is used.

Definition A **basis** is a linearly independent set of vectors that span the vector space. Let $\mathcal{B} = \left\{\boldsymbol{e}_{(\mu)}\right\} = \left\{\boldsymbol{e}_{(1)}, \cdots, \boldsymbol{e}_{(n)}\right\}$ denote a basis for $\mathcal{V}$. Vectors will be displayed in bold face. Parentheses will be used on subscripts to emphasize that $\boldsymbol{e}_{(\mu)}$ are vectors.

Definition Every vector in $\mathcal{V}$ can be uniquely expressed as a **linear combination** of basis vectors:

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

This expression uses the Einstein summation convention to add products of terms having matching upper and lower Greek indices. The lack of parentheses on the superscripts of V emphasizes that they are components and not vectors. Components will not be put in bold face. Also, we write a vector $\boldsymbol{e}_{(\mu)}$ using a subscript μ but we write a vector component V^{μ} using a superscript μ. We choose to do this so that we can write the vector as $\boldsymbol{e}_{(\mu)} = \begin{pmatrix} e^1_{(\mu)} \\ \vdots \\ e^n_{(\mu)} \end{pmatrix}$. Otherwise, we would need two superscripts, one in parentheses and one not.

Notation To represent the abstract vector space $\mathcal{V}$ as a vector space of column vectors we identify the basis vectors $\boldsymbol{e}_{(\mu)}$ as column vectors:

$$\boldsymbol{e}_{(\mu)} = \begin{pmatrix} 0 \\ \vdots \\ 1^{\mu} \\ \vdots \\ 0 \end{pmatrix}.$$

Every vector in $\mathcal{V}$ can be expressed as a linear combination of the basis vectors:

$$\boldsymbol{V} = V^{\mu}\boldsymbol{e}_{(\mu)} = V^{\mu}\begin{pmatrix} 0 \\ \vdots \\ 1^{\mu} \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} V^1 \\ \vdots \\ V^{\mu} \\ \vdots \\ V^n \end{pmatrix}.$$

The component V^{μ} is sometimes loosely referred to as a vector whereas a vector is more correctly expressed as $\boldsymbol{V} = V^{\mu}\,\mathbf{e}_{(\mu)}$. This is done because $\mathbf{e}_{(\mu)}$ are only place-holder zeros and ones; the information content is carried by the components.

Definition A **covector** is a row vector $\boldsymbol{\omega} = \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix}$. As with vectors, ω_{μ} is sometimes loosely referred to as a covector. Note that if $\boldsymbol{V}$ is a vector, then its **transpose**, $\boldsymbol{V}^{\mathsf{T}}$, is a covector.

Notation Let $\mathcal{V}^*$ be an n-dimensional vector space of covectors and $\mathcal{B}^* = \left\{\varepsilon^{(\nu)}\right\} = \left\{\varepsilon^{(1)},\cdots,\varepsilon^{(n)}\right\}$ a basis where $\varepsilon^{(\nu)} = \begin{pmatrix} 0 & \cdots & 1_{\nu} & \cdots & 0 \end{pmatrix}$. $\mathcal{V}^*$ is called a **covector space**. Every covector can be expressed as a linear combination $\boldsymbol{\omega} = \boldsymbol{\omega}_{\nu}\varepsilon^{(\nu)} = \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix}$ of basis covectors.

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

$$\boldsymbol{\omega}\,\boldsymbol{V} = \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix}\begin{pmatrix} V^1 \\ \vdots \\ V^n \end{pmatrix} = \omega_\mu V^\mu \in \mathbf{F}. \tag{i}$$

Definition $\boldsymbol{\omega}$ can be regarded as a **linear functional**:

$$\omega : \mathcal{V} \to \mathbf{F} : \omega(\boldsymbol{V}) = \boldsymbol{\omega}\boldsymbol{V}.$$

A **functional** is a function whose range is **F. Linear** means that the vector space operations of addition and scalar multiplication are preserved:

$$\omega(aV + b\boldsymbol{W}) = a\,\omega(\boldsymbol{V}) + b\,\omega(\boldsymbol{W}).$$

As illustrated, we use boldface for covectors like $\boldsymbol{\omega}$ but non-boldface for functions like ω.

Construction Since the linear functionals ω can be mapped 1-1 with the covectors $\boldsymbol{\omega}$, it should not be surprising that the functionals can be structured to form a vector space. We define addition of functions:

$$(\omega + \xi)(V) \equiv \omega(\boldsymbol{V}) + \xi(\boldsymbol{V})$$

and scalar multiplication:

$$(a\,\omega)(\boldsymbol{V}) = a\left[\omega(\boldsymbol{V})\right].$$

Then the function $\omega + \xi$ corresponds to the covector $\boldsymbol{\omega} + \boldsymbol{\xi}$ and the function $\alpha\omega$ corresponds to the covector $\alpha\boldsymbol{\omega}$. The additive identity is the zero function, $-\omega$ is the additive inverse of ω, and all the remaining vector space rules for addition and scalar multiplication are easily verified. We have proven the following theorem.

Theorem The mapping $\omega \mapsto \boldsymbol{\omega}$ is an **isomorphism**. That is, it is a 1-1 mapping that preserves the vector space structure of addition and scalar multiplication.

If we have a functional, ω, we are free to switch on-the-fly and consider it a covector, $\boldsymbol{\omega}$, and vice-versa.

Definition We have shown that $\mathcal{V}^*$ can thus be viewed from three perspectives: (1) abstract vector space, (2) vector space of functions, and (3) vector space of row vectors. In the functional perspective, $\mathcal{V}^*$ is called the **dual space of $\mathcal{V}$**. It is the vector space of linear transformations from $\mathcal{V}$ to **F**. In this perspective, the linear functionals ω are known as **dual vectors**. The abstract perspective simply means to view $\mathcal{V}^*$ as a set of abstract vector-space objects: $\boldsymbol{\omega} = \boldsymbol{\omega}_\nu \boldsymbol{\varepsilon}^{(\nu)}$.

Observe that in the matrix-algebra perspective we have

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

To be consistent, when we work in either the linear functional perspective or the abstract perspective, we will assume that the bases of $\mathcal{V}$ and $\mathcal{V}^*$ satisfy $\boldsymbol{\varepsilon}^{(\nu)} \mathbf{e}_{(\mu)} = \delta^\nu_\mu$.

In most books the terms "covector", "linear functional", "dual vector", and "1-form" are used interchangeably. There are actually subtle distinctions between them. For the record, we distinguish them now.

1. A **covector** refers to a row vector and belongs to the matrix-algebra perspective.
2. A **linear functional** is a linear map from a vector space to a field **F** and belongs to the linear functional perspective.
3. 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.** They are "*vectors*" because we are considering them as part of a vector space of linear functionals.
4. The term 1-form is most often used in differential geometry, which plays a role in general relativity. A **manifold** is a topological space that is locally homeomorphic to Euclidean space. A **vector field**, or just a **field**, is a set of vectors, one for every point on a manifold. A **1-form** is a vector field where the vectors are linear functionals. Somewhat confusingly, the term "1-form" is loosely and frequently used also to refer to individual linear functionals in the field. In physics, the term 1-form is most often used when the linear functionals are differentials, ∂f. In tensor theory, a 1-form refers to a (0, 1) tensor, soon to be defined, which is simply a dual vector.

Definition One can define $\mathcal{V}^{**}$, the dual space of the dual space $\mathcal{V}^*$**.** This space is isomorphic to $\mathcal{V}$ via the natural 1-1 mapping $V^{**}(\boldsymbol{\omega}) \equiv \omega(\boldsymbol{V})$, where $V^{**} \in \mathcal{V}$. In this way, $\mathcal{V}$ can be regarded as the dual space of $\mathcal{V}^*$, just as $\mathcal{V}^*$ is the dual space of $\mathcal{V}$.

A way to understand the importance of this is as follows. Just as a row vector can be viewed as a linear functional when it acts of a column vector, a column vector can be viewed as a linear functional when it acts on a row vector. Tensors are defined as linear functionals, so the perspective that row and column vectors are duals of each other plays a part.

We have learned how to go from row and column vectors to linear functionals. We also need to perform this process in reverse: identify a linear functional with a vector. Fortunately, this is very easy.

Construction Let $\left\{ T : \mathcal{V} \to \boldsymbol{F} \right\}$ be a vector space of linear functionals and let $\left\{ e_{(\mu)} \right\} = \left\{ e_{(1)}, \cdots, e_{(n)} \right\}$ be a basis, where each $e_{(\mu)} : \mathcal{V} \to \boldsymbol{F}$ is, of course, a linear functional. First, express T abstractly in terms of the basis elements: $T = T^{\mu} e_{(\mu)}$.

Next, define column vectors $\boldsymbol{e}_{(\mu)} = \begin{pmatrix} 0 \\ \vdots \\ 1^{\mu} \\ \vdots \\ 0 \end{pmatrix}$. The collection $\left\{ \boldsymbol{e}_{(\mu)} \right\}$ forms a basis for the vector space of length n column vectors: $\boldsymbol{T} = T^{\mu}\, \boldsymbol{e}_{(\mu)} = \begin{pmatrix} T^1 \\ \vdots \\ T^n \end{pmatrix}$. We are free to think of T as a linear functional or as $\boldsymbol{T}$, a column vector. We will be careful to use boldface for matrix objects and non-boldface for linear operators.

We have actually just proved that *every* n-dimensional vector space is isomorphic to the vector space of length n column vectors. This, in turn, proves a very important theorem.

Theorem All n-dimensional vector spaces are isomorphic.

In this sense, there is just a single vector space of each finite dimension. For example, the nxn matrices, the linear functionals, and $\mathbb{R}^n$ are not the same, but in terms of just their vector space properties, they are identical.

Definition The **product $\mathcal{V} \times \mathcal{W}$ of two vector spaces** is defined to be the vector space of pairs ($\boldsymbol{V}$, $\boldsymbol{W}$). Addition and scalar multiplication are carried out just like with (x, y) pairs in $\mathbb{R}^2$:

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

As we have seen, every linear functional can be considered to be a column vector. A linear functional defined on a product of vector spaces can also be considered to be a matrix. This identification is the critical step (9) in Section 4 where we develop rank (1, 1) tensors. We carry out that step now.

Construction Suppose $\mathcal{V}$ and $\mathcal{W}$ are *n*-dimensional vector spaces, $\mathcal{V}^*$ the dual space of $\mathcal{V}$, $\left\{ T : \mathcal{V}^* \times \mathcal{W} \rightarrow \boldsymbol{F} \right\}$ a vector space of linear functionals, and $\left\{ e^{\nu}_{\mu} \right\}$ a basis for {*T*}. Let $T = T^{\mu}_{\nu}\, e^{\nu}_{\mu}$. We can identify the function e^{ν}_{μ} with the matrix

$$\boldsymbol{e}^{(\nu)}_{(\mu)} \equiv \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1^{\mu}_{\nu} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}, \tag{ii}$$

and $\left\{ \boldsymbol{e}^{(\nu)}_{(\mu)} \right\}$ forms a basis for the vector space of *n*x*n* matrices. We can then also identify the linear functional *T* with the matrix $\boldsymbol{T} = T^{\mu}_{\nu}\, \boldsymbol{e}^{(\nu)}_{(\mu)}$. Note that the indices of 1^{μ}_{ν} match *T* rather than ***e***. This is because, as explained earlier, a vector $\boldsymbol{e}_{(\mu)}$ has components with superscripts μ.

Notice above that the coefficient $T^{\ \mu}_{\ \nu}$ is the same for both the matrix ***T*** and the function *T*. We will interchangeably refer to the abstract symbol $T^{\ \mu}_{\ \nu}$ as a linear functional or a matrix.

In the literature there is no consistent matrix algebra representation of T^{μ}_{ν}, $T_{\mu\nu}$, and $T^{\mu\nu}$. Often, all three are treated as matrices. We will treat only the first one as a matrix. We prefer to represent the middle one as a (1xn^2) row vector, and the 3rd one as a (n^2x1) column vector. Examples will be provided in Section 2. Because the vector spaces of (*n*x*n*) 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, treating all three of T^{μ}_{ν}, $T_{\mu\nu}$, and $T^{\mu\nu}$ as matrices is also legitimate.

It should be mentioned that linear operators that have 3 or more indices can be represented not only as row vectors, column vectors, and 2-dimensional matrices, but also 3-dimensional matrices, and more. We will treat all linear operators as either row vectors (e.g., $T_{\mu\nu o}$), column vectors (e.g., $T^{\mu\nu o\rho}$), or 2-D matrices (e.g., $T^{\mu\rho}_{\nu o\sigma}$).

2 Tensor Products of Matrices

Definition Equation (i), $\boldsymbol{\omega V} = \omega_\mu V^\mu$, is an example of an **inner product**, a sum of products of matching elements. Another kind of product is an **outer product**, where every element of $\boldsymbol{V}$ is multiplied by every element of ω:

$$\begin{pmatrix} V^1 \\ \vdots \\ V^n \end{pmatrix} \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix} = \begin{pmatrix} V^1\omega_1 & \cdots & V^1\omega_n \\ \vdots & & \vdots \\ V^n\omega_1 & \cdots & V^n\omega_n \end{pmatrix}$$

Definition Matrix multiplication is not defined to include outer products. Because it is useful for tensors, we introduce an outer product operation called **tensor product**, denoted by $\otimes$. We define $\boldsymbol{A} \otimes \boldsymbol{B}$ where $\boldsymbol{A}$ and $\boldsymbol{B}$ are any matrix objects: scalars, row vectors, column vectors, or matrices. We begin by defining $\otimes$ for a scalar k and a matrix object $\boldsymbol{A}$, where $\boldsymbol{A}$ is either a matrix, a row vector, or a column vector.

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

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

We complete the definition by defining $\otimes$ for two non-scalar matrix objects:

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

We refer to this as a *pattern definition* because the operation is easily carried out by following the pattern (iv).

Example 1: Two matrices:

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

$$\boldsymbol{A} \otimes \boldsymbol{B} \overset{\text{(iv)}}{=} \begin{pmatrix} a_{11} \otimes \boldsymbol{B} & a_{12} \otimes \boldsymbol{B} \\ a_{21} \otimes \boldsymbol{B} & a_{21} \otimes \boldsymbol{B} \end{pmatrix} = \begin{pmatrix} a_{11} \otimes b_{11} & a_{11} \otimes b_{12} & | & a_{12} \otimes b_{11} & a_{12} \otimes b_{12} \\ a_{11} \otimes b_{21} & a_{11} \otimes b_{22} & | & a_{12} \otimes b_{21} & a_{12} \otimes b_{22} \\ --- & --- & + & --- & --- \\ a_{21} \otimes b_{11} & a_{21} \otimes b_{12} & | & a_{22} \otimes b_{11} & a_{22} \otimes b_{12} \\ a_{21} \otimes b_{21} & a_{21} \otimes b_{22} & | & a_{22} \otimes b_{21} & a_{22} \otimes b_{22} \end{pmatrix}$$

Example 2: Column vector and a row vector:

$$\boldsymbol{V}\otimes\boldsymbol{\omega}\overset{\text{(iv)}}{=}\begin{pmatrix} V^1\otimes\boldsymbol{\omega} \\ \vdots \\ V^n\otimes\boldsymbol{\omega}\end{pmatrix}=\begin{pmatrix} V^1\otimes\begin{pmatrix}\omega_1 & \cdots & \omega_n\end{pmatrix} \\ \vdots \\ V^n\otimes\begin{pmatrix}\omega_1 & \cdots & \omega_n\end{pmatrix}\end{pmatrix}$$

$$\equiv\begin{pmatrix} V^1\otimes\omega_1 & \cdots & V^1\otimes\omega_n \\ \vdots & & \vdots \\ V^n\otimes\omega_1 & \cdots & V^n\otimes\omega_n\end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \left(V^1,\,\omega_1\right) & \cdots & \left(V^1,\,\omega_n\right) \\ \vdots & & \vdots \\ \left(V^n,\,\omega_1\right) & \cdots & \left(V^n,\,\omega_n\right)\end{pmatrix}$$

The tensor symbols and the ordered pairs inside the matrices are simply reminders not to combine the scalars V^μ and ω_μ.

As a particular example of this, we can re-generate equation (ii) as the tensor product of a column basis vector and a row basis vector:

$$\boldsymbol{e}_{(\mu)}\otimes\boldsymbol{\varepsilon}^{(\nu)}=\begin{pmatrix}\vdots \\ 1^\mu \\ \vdots\end{pmatrix}\otimes\begin{pmatrix}\cdots & 1_\nu & \cdots\end{pmatrix}=\begin{pmatrix} & \vdots & \\ \cdots & 1^\mu_\nu & \cdots \\ & \vdots & \end{pmatrix} \tag{v}$$

Example 3: Two row vectors:

$$\begin{pmatrix}\omega_1 & \cdots & \omega_n\end{pmatrix}\otimes\begin{pmatrix}\omega_1 & \cdots & \omega_n\end{pmatrix}\overset{\text{(iv)}}{=}\begin{pmatrix}\omega_1\begin{pmatrix}\omega_1 & \cdots & \omega_n\end{pmatrix} & \cdots & \omega_n\begin{pmatrix}\omega_1 & \cdots & \omega_n\end{pmatrix}\end{pmatrix}$$

$$\begin{pmatrix}\omega_1\omega_1 & \cdots & \omega_1\omega_n & \omega_2\omega_1 & \cdots & \omega_2\omega_n & \cdots & \omega_n\omega_1 & \cdots & \omega_n\omega_n\end{pmatrix}$$

Whereas $\boldsymbol{\omega}$ is an *n* covector, $\boldsymbol{\omega}\otimes\boldsymbol{\omega}$ is an n^2 covector, just as $\boldsymbol{V}\otimes\boldsymbol{\omega}$, above, is an *nxn* matrix. Similarly, the tensor product of two length *n* column vectors would be expressed as a length n^2 column vector.

3 Tensors

We now have the machinery to define tensors. Though tensors can be defined for infinite-dimensional vector spaces, we will restrict our scope to finite-dimensional vector spaces. Prior to step 9 we consider tensors strictly as functions. After step 9, we also view them as matrix algebra objects.

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

$$T : \underbrace{\Upsilon^* \times \cdots \times \mathcal{Z}^*}_{k \text{ terms}} \times \underbrace{\mathcal{V} \times \cdots \times \mathcal{W}}_{\ell \text{ terms}} \rightarrow \mathbf{F}. \tag{1}$$

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

$$\begin{aligned}&T\left(a_1 \boldsymbol{\omega}^{(1)} + a_2 \boldsymbol{\omega}^{(2)}, b_1 \boldsymbol{V}_{(1)} + b_{(2)} \boldsymbol{V}_{(2)}\right)\\&= a_1 b_1 T\left(\boldsymbol{\omega}^{(1)}, \boldsymbol{V}_{(1)}\right) + a_1 b_2 T\left(\boldsymbol{\omega}^{(1)}, \boldsymbol{V}_{(2)}\right) + a_2 b_1 T\left(\boldsymbol{\omega}^{(2)}, \boldsymbol{V}_{(1)}\right) + a_2 b_2 T\left(\boldsymbol{\omega}^{(2)}, \boldsymbol{V}_{(2)}\right).\end{aligned}$$

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

As with linear functionals, addition and scalar multiplication of tensor functions are defined naturally, resulting that the *set* of (k, ℓ) tensors form a *vector space* under addition and scalar multiplication. In particular,

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

Just as a linear functional T can be identified with the matrix $\boldsymbol{T} = T^{\mu}_{\ \nu}\, \mathbf{e}^{(\nu)}_{(\mu)}$, we will show that a tensor is isomorphic to a matrix. (If k = 0, it is isomorphic to a row vector. If ℓ = 0, it is isomorphic to a column vector.)

To keep development simple, we begin with the (1, 1) tensors.

4 Rank (1,1) Tensors

Tensors and their properties are developed in 17 steps, summarized below.

(1) $\mathcal{V}\otimes\mathcal{W}^* \equiv \left\{T : \mathcal{V}^* \times \mathcal{W} \rightarrow \mathbf{F} : \ T \text{ is bilinear}\right\}$

(2) $\mathcal{B}_{\mathcal{V}} = \left\{\mathbf{e}_{(\mu)}\right\}_{\mu=1}^{n}$, $\mathcal{B}_{\mathcal{W}} = \left\{\boldsymbol{f}_{(\rho)}\right\}_{\rho=1}^{m}$, $\mathcal{B}_{\mathcal{V}^*} = \left\{\varepsilon^{(\nu)}\right\}_{\nu=1}^{n}$, and $\mathcal{B}_{\mathcal{W}^*} = \left\{\varphi^{(\sigma)}\right\}_{\sigma=1}^{m}$

(3) $\boldsymbol{V} = V^{\mu}\,\mathbf{e}_{(\mu)} \in \mathcal{V}$, $\boldsymbol{W} = W^{\rho}\,\boldsymbol{f}_{(\rho)} \in \mathcal{W}$, $\boldsymbol{\omega} = \omega_{\nu}\,\varepsilon^{(\nu)} \in \mathcal{V}^*$, and $\boldsymbol{\xi} = \xi_{\sigma}\,\varphi^{(\sigma)} \in \mathcal{W}^*$

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

(5) $e^{\nu}_{\mu} : \mathcal{V}^* \times \mathcal{W} : e^{\nu}_{\mu}\left(\boldsymbol{\omega}, \boldsymbol{W}\right) = \omega_{\mu}\,W^{\nu}$

(6) $\mathcal{B}_{\mathcal{V}\otimes\mathcal{W}^*} = \left\{e^{\nu}_{\mu}\right\}_{\mu=1}^{n}\,{}_{\nu=1}^{m}$

(7) $T^{\mu}_{\nu} \equiv T\left(\varepsilon^{(\mu)}, \boldsymbol{f}_{(\nu)}\right)$

(8) $T = T^{\mu}_{\nu}\,e^{\nu}_{\mu}$

(9) There is a 1-1 map between the functionals e^{ν}_{μ} and the matrices $\mathbf{e}^{(\nu)}_{(\mu)}$

(10) $\boldsymbol{V}\otimes\boldsymbol{\xi} \equiv V^{\mu}\,\xi_{\nu}\ \mathbf{e}^{(\nu)}_{(\mu)}$

(11) $\boldsymbol{e}_{(\mu)} \otimes \varphi^{(\nu)} = \mathbf{e}^{(\boldsymbol{\nu})}_{(\boldsymbol{\mu})}$

(12) $\boldsymbol{T} = T^{\mu}_{\nu}\,\mathbf{e}_{(\mu)} \otimes \varphi^{(\nu)}$

(13) $\mathcal{V}\otimes\mathcal{W}^* = \left\{T^{\mu}_{\nu}\,\mathbf{e}_{(\mu)} \otimes \varphi^{(\nu)}\right\}$

(14) $T\left(\boldsymbol{\omega}, \boldsymbol{W}\right) = T^{\mu}_{\nu}\,\omega_{\mu}\,W^{\nu}$

(15) $T\left(\boldsymbol{\omega}, \boldsymbol{W}\right) = \boldsymbol{\omega}\,\boldsymbol{T}\,\boldsymbol{W}$

(16) $\boldsymbol{T} = \boldsymbol{T}_{\boldsymbol{\omega}} \otimes \boldsymbol{T}^{\boldsymbol{W}}$

(17) If $\boldsymbol{R} = R^{\mu}_{\nu}\,\mathbf{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{(\nu)} \in \mathcal{V}\otimes\mathcal{W}^*$ and $\boldsymbol{S} = S^{\sigma}_{\rho}\,\boldsymbol{f}_{(\sigma)} \otimes \varepsilon^{(\rho)} \in \mathcal{W}\otimes\mathcal{V}^*$, define a (2*k*, 2*ℓ*) tensor in $\mathcal{V}\otimes\mathcal{W}\otimes\mathcal{W}^*\otimes\mathcal{V}^*$ by

$$\boldsymbol{T} = \boldsymbol{R}\otimes\boldsymbol{S} = R^{\mu}_{\nu}\,\mathbf{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{(\nu)} \otimes\ S^{\sigma}_{\rho}\,\boldsymbol{f}_{(\sigma)} \otimes \varepsilon^{(\rho)}$$

$$\equiv R^{\mu}_{\nu}\ S^{\sigma}_{\rho}\,\mathbf{e}_{(\mu)} \otimes \boldsymbol{f}_{(\sigma)} \otimes \boldsymbol{\varphi}^{(\nu)} \otimes \varepsilon^{(\rho)} = T^{\mu\sigma}_{\nu\rho}\,\mathbf{e}_{(\mu)} \otimes \boldsymbol{f}_{(\sigma)} \otimes \boldsymbol{\varphi}^{(\nu)} \otimes \boldsymbol{\varepsilon}^{(\rho)}$$

where $T^{\mu\sigma}_{\nu\rho} = R^{\mu}_{\nu}\,S^{\sigma}_{\rho}$.

First, we give an overview of the steps and then we provide derivation of the equations and proof that the product basis in (6) is indeed a basis.

Overview of the Steps

(1) Definition of tensor product of a covector space and a vector space in terms of linear functionals. This is the same equation (1) given in Section 3 but tailored for just one covector space and one vector space.

Note 1. Pay particular attention to $\mathcal{V}\otimes\mathcal{W}^*$ on LHS and $\mathcal{V}^*\otimes\mathcal{W}$ on RHS. This occurs because, as we learned in Section 1, a *vector* acts as a function on a *covecto*r, and a *covector* acts as a function on a *vecto*r. LHS consists of vectors and covectors. RHS consists of functions.

Note 2. In Section 2 we defined $\otimes$ for matrices. Matrices will not be introduced until step (9). Until then, T should be regarded only as a linear functional. Moreover, in (16) we will explicitly express T as a tensor product of vectors, $\boldsymbol{T} = \boldsymbol{V} \otimes \boldsymbol{W}$. When that happens, it will justify use of the tensor symbol in the vector space name on the LHS of (1). Until then, $\mathcal{V}\otimes\mathcal{W}$, too**,** should be regarded as nothing more than a vector space of functions, unrelated to tensor operations.

(2–3) Abstract representation of vectors and covectors in terms of bases. Because there are only two spaces, it is easy enough for now to allow them different dimensions. When we return to $k + \ell$ spaces, we will give them all the same dimension n to keep the notation from getting overly complex.

(4–8) This section defines a collection of product linear functionals and then proves that they form a basis for $V \otimes W^*$. Step (8) shows how to express a linear functional T in terms of basis elements.

(9–13) The tensor operation has been defined in Section 2 for matrices. This section extends tensors to include linear functionals. Step (9) is the construction of the basis matrix that corresponds to the functional basis. This was carried out in Section 1 as equation (ii). Equation (10) defines the tensor product between a vector and covector in terms of the matrix. Equation (11) shows the resulting tensor product between 2 basis elements. Equation (12) expresses a linear functional as a tensor product, justifying that these linear functionals are tensors. The coefficients, T^{μ}_{ν}, were identified in Section 1 to be nxn matrices, so from the matrix algebra perspective a rank (1, 1) tensor is a matrix.

(14–16) Equation (14) at last provides a formula for $T(\omega, \boldsymbol{W})$. It is an abstract vector space representation. Equation (15) provides the equivalent matrix-algebra expression for $T(\omega, \boldsymbol{W})$. Equation (16) states that the rank (1, 1) tensor matrix can also be viewed as a tensor product of a column vector and a row vector. This can also be seen in (11).

(17) The formula for the tensor product of two (1, 1) tensors.

Derivation

Equations (1–4) are definitions.

Equation (5) extends the definition (4) of e_μ^ν from basis elements to covector-vector pairs:

$$e_\mu^\nu : \mathcal{V}^* \times \mathcal{W} : e_\mu^\nu\left(\boldsymbol{\omega}, \boldsymbol{W}\right) = \omega_\mu W^\nu : \tag{5}$$

$$e_\mu^\nu\left(\boldsymbol{\omega}, \boldsymbol{W}\right) \overset{(3)}{=} e_\mu^\nu\left(\omega_\sigma \boldsymbol{\varepsilon}^{(\sigma)}, W^\rho \boldsymbol{f}_{(\rho)}\right) \overset{(4)}{=} \omega_\sigma W^\rho \; \delta_\mu^\sigma \; \delta_\rho^\nu = \omega_\mu W^\nu \qquad \checkmark$$

In defining e_ν^μ, there are reasons why some books slant the indices NW to SE or SW to NE:

1. Without slanting, one cannot immediately tell in (5) which index operates on $\boldsymbol{\omega}$ and which operates on $\boldsymbol{W}$. In fact, μ acts on $\boldsymbol{\omega}$ and ν acts on $\boldsymbol{W}$. Correct slanting would be from SW to NE.

2. If an operation involves, for example, a tensor transpose, the slant could identify the tensor versus its transpose. Simply reversing indices to indicate the transpose could be confusing since the indices are dummies and we are free to interchange them at any time.

Nonetheless, we have chosen to keep indices vertical because the slant takes up a lot of space when we get to multiple superscripts and subscripts.

To prove that $e_\mu^\nu \in \mathcal{V} \otimes \mathcal{W}^*$, we must show that e_μ^ν is bilinear:

Suppose $a\boldsymbol{\omega} + b\boldsymbol{\xi} \overset{(3)}{=} \left(a\omega_\nu + b\xi_\nu\right)\boldsymbol{\varepsilon}^{(\nu)} \equiv \upsilon_\nu \boldsymbol{\varepsilon}^{(\nu)}$ and

$c\boldsymbol{W} + d\boldsymbol{V} \overset{(3)}{=} \left(cW^\rho + dV^\rho\right)\boldsymbol{f}_{(\rho)} \equiv U^\rho \boldsymbol{f}_{(\rho)}$. Then

$$e_\mu^\nu\left(a\boldsymbol{\omega} + b\boldsymbol{\xi}, c\boldsymbol{W} + d\boldsymbol{V}\right) = e_\mu^\nu\left(\upsilon_\nu \boldsymbol{\varepsilon}^{(\nu)},\; U^\rho \boldsymbol{f}_{(\rho)}\right) \overset{(5)}{=} \upsilon_\mu U^\nu$$

$$= \left(a\omega_\mu + b\xi_\mu\right)\left(cW^\nu + dV^\nu\right) = ac\,\omega_\mu W^\nu + ad\,\omega_\mu V^\nu + bc\,\xi_\mu W^\nu + bd\,\xi_\mu V^\nu$$

$$= ac\,e_\mu^\nu\left(\boldsymbol{\omega}, \boldsymbol{W}\right) + ad\,e_\mu^\nu\left(\boldsymbol{\omega}, \boldsymbol{V}\right) + bc\,e_\mu^\nu\left(\boldsymbol{\xi}, \boldsymbol{W}\right) + bd\,e_\mu^\nu\left(\boldsymbol{\xi}, \boldsymbol{V}\right). \qquad \checkmark$$

Define $\mathcal{B}_{\mathcal{V}\otimes\mathcal{W}^*}$:

$$\mathcal{B}_{\mathcal{V}\otimes\mathcal{W}^*} = \left\{e_\mu^\nu\right\}_{\mu=1}^{n}{}_{\nu=1}^{m} \tag{6}$$

In order to show that it is a basis, we must show that it spans $\mathcal{V} \otimes \mathcal{W}^*$ and is linearly independent.

$\mathcal{B}_{\mathcal{V}\otimes\mathcal{W}^*}$ spans $\mathcal{V}\otimes\mathcal{W}^*$:

For $T\in\mathcal{V}\otimes\mathcal{W}^*$, define

$$T^{\mu}_{\ \nu} \equiv T\left(\varepsilon^{(\mu)}, \boldsymbol{f}_{(\nu)}\right). \tag{7}$$

Let $\boldsymbol{\omega} \stackrel{(3)}{=} \omega_{\mu}\,\varepsilon^{(\mu)} \in \mathcal{V}^*$ and $\boldsymbol{W} = W^{\nu}\boldsymbol{f}_{(\nu)} \in \mathcal{W}^*$. Then

$$T\left(\boldsymbol{\omega},\boldsymbol{W}\right) = T\left(\omega_{\mu}\,\varepsilon^{(\mu)},\ W^{\nu}\boldsymbol{f}_{(\nu)}\right) \stackrel{\text{(bilinear)}}{=} \omega_{\mu}W^{\nu}\,T\left(\varepsilon^{(\mu)},\ \boldsymbol{f}_{(\nu)}\right) \stackrel{(6)}{=} T^{\mu}_{\ \nu}\,\omega_{\mu}W^{\nu}$$

$$\stackrel{(5)}{=} T^{\mu}_{\ \nu}\,e^{\nu}_{\ \mu}\left(\boldsymbol{\omega},\boldsymbol{W}\right)$$

$$\Rightarrow\quad T = T^{\mu}_{\ \nu}\,e^{\nu}_{\ \mu} \quad ✔ \tag{8}$$

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

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

Note that we have proven in passing that dim($\mathcal{V}\otimes\mathcal{W}^*$) = nm since that is the size of its basis.

We are at last in a position to define the tensor product. Until now everything has been about linear functionals even though we have used the tensor symbol in some of the names.

$$\text{Let}\quad \mathbf{e}^{(\nu)}_{(\mu)} = \begin{pmatrix} & \vdots & \\ \cdots & 1^{\mu}_{\ \nu} & \cdots \\ & \vdots & \end{pmatrix} \tag{9}$$

be the *nxm* matrix that corresponds to the linear functional $e^{\nu}_{\ \mu}$, as constructed in Section (1) and presented in Equation (ii).

Define the tensor product of a vector $\boldsymbol{V}$ and covector ξ to be the matrix

$$\boldsymbol{V}\otimes\boldsymbol{\xi} \equiv V^{\mu}\,\xi_{\nu}\ \mathbf{e}^{(\nu)}_{(\mu)}. \tag{10}$$

Recall that $\boldsymbol{V}$ and ξ were defined in (3). Matrix (10) was worked out in Example 2 of Section 2.

Claim: $\boldsymbol{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{\nu} = \boldsymbol{e}_{(\mu)}^{(\nu)}$ (11)

$\boldsymbol{e}_{(\mu)} = \delta^{\alpha}_{\mu}\, \boldsymbol{e}_{(\alpha)}$. Let $\boldsymbol{V} = \boldsymbol{e}_{(\mu)}$. Then $\boldsymbol{V} = V^{\alpha}\, \boldsymbol{e}_{(\alpha)}$. So, $V^{\alpha} = \delta^{\alpha}_{\mu}$. Similarly, let $\boldsymbol{\xi} = \boldsymbol{\varphi}^{(\nu)}$. Then $\boldsymbol{\varphi}^{(\nu)} = \delta^{\nu}_{\beta}\, \boldsymbol{\varphi}^{(\beta)}$, $\boldsymbol{\xi} = \xi_{\beta}\, \boldsymbol{\varphi}^{(\beta)}$, and $\xi_{\beta} = \delta^{\nu}_{\beta}$. Thus,

$$\boldsymbol{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{(\nu)} = \boldsymbol{V} \otimes \boldsymbol{\xi} \overset{(9)}{=\!=} V^{\alpha}\xi_{\beta}\, \boldsymbol{e}_{(\alpha)}^{(\beta)} = \delta^{\alpha}_{\mu}\, \delta^{\nu}_{\beta}\, \boldsymbol{e}_{(\alpha)}^{(\beta)} = \boldsymbol{e}_{(\mu)}^{(\nu)} \quad \checkmark$$

This claim may easier to visualize in matrix form, shown in Equation (v) of Section 2.

Combining (11) with $\boldsymbol{T} = T^{\mu}_{\nu}\, \boldsymbol{e}_{(\mu)}^{(\nu)}$, the matrix form of (8), yields

$$\boldsymbol{T} = T^{\mu}_{\nu}\, \boldsymbol{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{(\nu)} \tag{12}$$

Combining (12) with (1) gives

$$\mathcal{V} \otimes \mathcal{W}^{*} = \left\{ T^{\mu}_{\nu}\, \boldsymbol{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{(\nu)} \right\} \tag{13}$$

This justifies the $\mathcal{V} \otimes \mathcal{W}$ designation in (1). We can now explicitly view $\mathcal{V} \otimes \mathcal{W}$ as a set of tensor products. Note that a tensor T^{μ}_{ν}, which acts on the product of two covectors, is a tensor product of vectors, not of covectors.

From (12): A (1, 1) tensor is a tensor product of a vector with a covector. From (13), $\mathcal{V} \otimes \mathcal{W}^{*}$ is isomorphic to the n^2 dimensional vector space of tensors T^{μ}_{ν}. We know from Section 2 that the tensors T^{μ}_{ν} are *n*x*n* matrices.

One more observation:

$$T(\omega, \boldsymbol{W}) = T^{\mu}_{\nu}\, \omega_{\mu} W^{\nu} \tag{14}$$

$$T(\omega, \boldsymbol{W}) \overset{(3)}{=} T\left(\omega_{\mu}\varepsilon^{(\mu)}, W^{\nu}\boldsymbol{f}_{(\nu)}\right) \overset{\text{(bilinearity)}}{=} \omega_{\mu} W^{\nu}\, T\left(\varepsilon^{(\mu)}, \boldsymbol{f}_{(\nu)}\right) \overset{(7)}{=} T^{\mu}_{\nu}\, \omega_{\mu} W^{\nu}.$$

Because RHS of (14) is a sum of products, the order of the terms does not matter. But, if we were to express RHS of (14) in matrix-algebra terminology, the order very much matters. In what order would we epxress the product:

$\boldsymbol{T}\omega\boldsymbol{W}$, $\omega\boldsymbol{W}\boldsymbol{T}$, $\omega\boldsymbol{T}\boldsymbol{W}$, or maybe $\boldsymbol{W}\omega\boldsymbol{T}$?

Since the result must be a scalar, the order is clear. Because ω is a row vector, there must be a column vector, $\boldsymbol{T}_{\omega}$, to its right. Because $\boldsymbol{W}$ is a column vector, there must be a row vector, $\boldsymbol{T}^{\boldsymbol{W}}$, to its left. So, we must have that

$$\boldsymbol{T} = \boldsymbol{T}_\omega \boldsymbol{T}^{\boldsymbol{W}}$$

and

$$T\left(\omega, \boldsymbol{W}\right) = \left(\omega\, \boldsymbol{T}_\omega\right)\left(\boldsymbol{T}^{\boldsymbol{W}} \boldsymbol{W}\right) = \omega\, \boldsymbol{T} \boldsymbol{W} \; .$$

This works in so far as $\omega \boldsymbol{T}_\omega$ and $\boldsymbol{T}^{\boldsymbol{W}} \boldsymbol{W}$ are scalars. However, it doesn't work in so far as matrix multiplication is not defined between a column vector $\boldsymbol{T}_\omega$ and a row vector $\boldsymbol{T}^{\boldsymbol{W}}$. In Section 2 we showed that this must be a tensor product. So, the matrix algebra version of equation (14) becomes

$$T\left(\omega, \boldsymbol{W}\right) = \omega \boldsymbol{T}\, \boldsymbol{W} \tag{15}$$

where

$$\boldsymbol{T} = \boldsymbol{T}_\omega \otimes \boldsymbol{T}^{\boldsymbol{W}} \; . \tag{16}$$

Observe that while the location of ω, $\boldsymbol{W}$, and $\boldsymbol{T}$ matters, the order of execution does not. One can compute $\omega \boldsymbol{T}_\omega$ first, or $\boldsymbol{T}_\omega \otimes \boldsymbol{T}^{\boldsymbol{W}}$ first, or $\boldsymbol{T}^{\boldsymbol{W}} \boldsymbol{W}$ first.

Although a tensor T has been defined as a linear function acting on a pair $\left(\omega,\ \boldsymbol{W}\right)$, there is nothing forcing it to act on both. For example, we can think of T as a linear function from $\mathcal{W}$ to $\mathcal{W}$ as follows. Since $\boldsymbol{T}^{\boldsymbol{W}} \boldsymbol{W}$ is a scalar, $\boldsymbol{T}\boldsymbol{W} = \boldsymbol{T}_\omega \boldsymbol{T}^{\boldsymbol{W}} \boldsymbol{W} = \left(\boldsymbol{T}^{\boldsymbol{W}} \boldsymbol{W}\right)\boldsymbol{T}_\omega$ is a column vector. So $T\,|_{\mathcal{W}}\colon \mathcal{W} \to \mathcal{W} : T\left(\boldsymbol{W}\right) = \boldsymbol{T}\boldsymbol{W}$ and $T\left(\boldsymbol{W}\right) \in \mathcal{W}$. Similarly we can think of T as a linear function from $\mathcal{V}^*$ to $\mathcal{V}^*$ since $\omega \boldsymbol{T} \in \mathcal{V}^*$

Not only can a tensor act on vectors and covectors, a tensor can also act on another tensor to produce a tensor. Step (17) defines a special case of how to take the tensor product of tensors when the tensors are defined over different domains.

If $\boldsymbol{R} = R^\mu_{\ \nu}\, \boldsymbol{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{(\nu)} \in \mathcal{V} \otimes \mathcal{W}^*$ and $\boldsymbol{S} = S^\sigma_{\ \rho}\, \boldsymbol{f}_{(\sigma)} \otimes \boldsymbol{\varepsilon}^{(\rho)} \in \mathcal{W} \otimes \mathcal{V}^*$ are tensors, define a ($2k$, 2ℓ) tensor in $\mathcal{V} \otimes \mathcal{W} \otimes \mathcal{W}^* \otimes \mathcal{V}^*$ by

$$\begin{aligned} \boldsymbol{T} = \boldsymbol{R} \otimes \boldsymbol{S} &= R^\mu_{\ \nu}\, \boldsymbol{e}_{(\mu)} \otimes \boldsymbol{\varphi}^{(\nu)} \otimes\ S^\sigma_{\ \rho}\, \boldsymbol{f}_{(\sigma)} \otimes \boldsymbol{\varepsilon}^{(\rho)} \\ &\equiv R^\mu_{\ \nu}\ S^\sigma_{\ \rho}\, \boldsymbol{e}_{(\mu)} \otimes \boldsymbol{f}_{(\sigma)} \otimes \boldsymbol{\varphi}^{(\nu)} \otimes \boldsymbol{\varepsilon}^{(\rho)} = T^{\mu\sigma}_{\ \ \nu\rho}\, \boldsymbol{e}_{(\mu)} \otimes \boldsymbol{f}_{(\sigma)} \otimes \boldsymbol{\varphi}^{(\nu)} \otimes \boldsymbol{\varepsilon}^{(\rho)} \end{aligned} \tag{17}$$

where $T^{\mu\sigma}_{\ \ \nu\rho} = R^\mu_{\ \nu} S^\sigma_{\ \rho}$.

We can view this as the tensor $\boldsymbol{R}$ acting on the tensor $\boldsymbol{S}$ to produce the tensor $\boldsymbol{T}$.

The current formula is a special case, a gentle introduction to the more general formula (17) in Section 6. The equation (17) above can be used, for example, to compute $\boldsymbol{R} \otimes \boldsymbol{S}(\omega, \xi, \boldsymbol{W}, \boldsymbol{V})$:

$$\boldsymbol{R} \otimes \boldsymbol{S}(\boldsymbol{\omega}, \xi, \boldsymbol{W}, \boldsymbol{V}) = \boldsymbol{T}(\boldsymbol{\omega}, \ \xi, \boldsymbol{W}, \boldsymbol{V}) \overset{(14)}{=} T^{\mu\sigma}_{\nu\rho} \, \omega_\mu \, \xi_\sigma W^\nu V^\rho \overset{(17)}{=} R^\mu_\nu S^\sigma_\rho \, \omega_\mu \, \xi_\sigma W^\nu V^\rho \, .$$

5 Lower Rank Tensors

A (2, 0) tensor space is $\mathcal{V} \otimes \mathcal{W} = \{T : \mathcal{V}^* \times \mathcal{W}^* \to \mathbf{F} : \ T \text{ is bilinear}\}$.

A (2, 0) tensor is formed from a tensor product of an *n* vector with an *m* vector. The result is that a (2, 0) tensor is an *nm* vector. As noted earlier, some books may treat a (2, 0) tensor as an *n*x*m* matrix.

A (0, 2) tensor space is $\mathcal{V}^* \otimes \mathcal{W}^* = \{T : \mathcal{V} \times \mathcal{W} \to \mathbf{F} : \ T \text{ is bilinear}\}$.

A (0, 2) tensor is an *nm* covector.

A (1, 0) tensor space is $\mathcal{V} = \{T : \mathcal{V}^* \to \mathbf{F} : \ T \text{ is linear}\}$.

A (1, 0) tensor is a vector because there is no product involved.

A (0, 1) tensor space is $\mathcal{V}^* = \{T : \mathcal{V} \to \mathbf{F} : \ T \text{ is linear}\}$

A (0, 1) tensor is a covector.

For a (0, 0) tensor, there is no product at all.
A (0, 0) tensor is a scalar.

6 Rank (*k*, ℓ) Tensors

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

We develop rank (k, ℓ) tensors following the steps (1–17) used for rank (1, 1). Since all of the proofs are straight-forward extensions of the proofs given in Section 4, they are omitted. We begin with equation (1), definition of a tensor.

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

To simplify this, we set

$$\mathcal{V} = \mathcal{V}_1 \times \cdots \times \mathcal{V}_k ,$$

$$\mathcal{W} = \mathcal{W}_1 \times \cdots \times \mathcal{W}_\ell ,$$

and

$$\mathcal{T} = \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_k \otimes \mathcal{W}_1{}^* \otimes \cdots \otimes \mathcal{W}_\ell{}^*$$

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

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

Since there is only one reference vector space $\mathcal{V}$, equations (2) and (3) can be expressed simply. We specify just a single basis for all the $\mathcal{V}_i$ and a single basis for all the $\mathcal{W}_j$:

$$\mathcal{B} = \left\{ \mathbf{e}_{(\mu)} \right\}_{\mu=1}^{n} \text{ and } \mathcal{B}^* = \left\{ \varepsilon^{(\nu)} \right\}_{\nu=1}^{n} \tag{2}$$

$$\begin{aligned}&\boldsymbol{V} = V^\mu \mathbf{e}_{(\mu)} \in \mathcal{V}, \ \ \boldsymbol{W} = W^\rho \mathbf{e}_{(\rho)} \in \mathcal{W}, \\ &\boldsymbol{\omega} = \omega_\nu \varepsilon^{(\nu)} \in \mathcal{V}^*, \text{ and } \boldsymbol{\xi} = \xi_\sigma \varepsilon^{(\sigma)} \in \mathcal{W}^*\end{aligned} \tag{3}$$

We define the linear functionals on the basis elements in (4) and then extend them to all members of $\mathcal{V}^* \times \mathcal{W}$ in (5):

$$e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}\left(\varepsilon^{(\sigma_1)}, \cdots, \varepsilon^{(\sigma_k)}, \mathbf{e}_{(\rho_1)}, \cdots, \mathbf{e}_{(\rho_\ell)} \right) \equiv \delta^{\sigma_1}_{\mu_1} \cdots \delta^{\sigma_k}_{\mu_k} \, \delta^{\nu_1}_{\rho_1} \cdots \delta^{\nu_\ell}_{\rho_\ell}. \tag{4}$$

$$e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k} : \mathcal{V}^* \times \mathcal{W} \to \mathbf{F} : e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}\left(\boldsymbol{\omega}^{(1)}, \cdots, \boldsymbol{\omega}^{(k)}, \boldsymbol{W}_{(1)}, \cdots, \boldsymbol{W}_{(\ell)} \right) = \omega_{\mu_1} \cdots \omega_{\mu_k} W^{\nu_1} \cdots W^{\nu_\ell} \tag{5}$$

We define the basis for $\mathcal{T}$ and expressions for the components.

$$\mathcal{B}_{\mathcal{T}} = \left\{ e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k} : \mu_1 = 1, \cdots, n; \cdots;\ \mu_k = 1, \cdots, n;\ \nu_1 = 1, \cdots, n; \cdots;\ \nu_\ell = 1, \cdots, n \right\} \tag{6}$$

$$T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} = T\left(\varepsilon^{(\mu_1)}, \cdots, \varepsilon^{(\mu_k)}, \mathbf{e}_{(\nu_1)}, \cdots, \mathbf{e}_{(\nu_\ell)} \right). \tag{7}$$

$$T = T^{\mu_1 \cdots \mu_\ell}_{\nu_1 \cdots \nu_k}\ e^{\nu_1 \cdots \nu_k}_{\mu_1 \cdots \mu_\ell} \tag{8}$$

Since the basis has size $n^{k+\ell}$, dim($\mathcal{T}$) = $n^{k+\ell}$.

There is a 1-1 map between the functionals $e^{\nu_1 \cdots \nu_\ell}_{\mu_1 \cdots \mu_k}$ and the matrices $\mathbf{e}^{(\nu_1) \cdots (\nu_\ell)}_{(\mu_1) \cdots (\mu_k)}$ (9)

Each $\mathbf{e}^{(\nu_1) \cdots (\nu_\ell)}_{(\mu_1) \cdots (\mu_k)}$ is an n^k x n^ℓ matrix having all zeros except for a "1" in the cell (n^k, n^ℓ). There are $n^{k+\ell}$ basis matrices $\mathbf{e}^{(\nu_1) \cdots (\nu_\ell)}_{(\mu_1) \cdots (\mu_k)}$. The parentheses on the indices indicate that we are enumerating matrices and not matrix components.

Define the tensor product of vectors and covectors:

$$\boldsymbol{V}_{(1)} \otimes \cdots \otimes \boldsymbol{V}_{(k)} \otimes \boldsymbol{\xi}^{(1)} \otimes \cdots \otimes \boldsymbol{\xi}^{(\ell)} \equiv V^{\mu_1} \cdots V^{\mu_k}\, \xi_{\nu_1} \cdots \xi_{\nu_\ell}\, \mathbf{e}^{(\nu_1) \cdots (\nu_\ell)}_{(\mu_1) \cdots (\mu_k)}. \tag{10}$$

As shown for (1, 1) tensors, this leads to the tensor product of basis vectors and covectors:

$$\mathbf{e}_{(\mu_1)} \otimes \cdots \otimes \mathbf{e}_{(\mu_k)} \otimes \varepsilon^{(\nu_1)} \otimes \cdots \otimes \varepsilon^{(\nu_\ell)} = \mathbf{e}^{(\nu_1) \cdots (\nu_\ell)}_{(\mu_1) \cdots (\mu_k)}: \tag{11}$$

Substituting (11) into (8) yields

$$T = T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} \mathbf{e}_{(\mu_1)} \otimes \cdots \otimes \mathbf{e}_{(\mu_k)} \otimes \varepsilon^{(\nu_1)} \otimes \cdots \otimes \varepsilon^{(\nu_\ell)} \tag{12}$$

and substituting (12) into (1') yields

$$\mathcal{T} = \left\{ T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} \mathbf{e}_{(\mu_1)} \otimes \cdots \otimes \mathbf{e}_{(\mu_k)} \otimes \varepsilon^{(\nu_1)} \otimes \cdots \otimes \varepsilon^{(\nu_\ell)} \right\}. \tag{13}$$

From (12), a (k, ℓ) tensor is a tensor product of k vectors and ℓ covectors. It is a n^kx n^ℓ matrix whose elements are the components $T^{\mu_1 \cdots \mu_\ell}_{\nu_1 \cdots \nu_k}$.

From (13), $\mathcal{T} = \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_k \otimes \mathcal{W}_1{}^* \otimes \cdots \otimes \mathcal{W}_\ell{}^*$ is isomorphic to the $n^{(k+\ell)}$-dimensional vector space of matrices $\boldsymbol{T} = T^{\mu_1 \cdots \mu_\ell}_{\nu_1 \cdots \nu_k}\ \mathbf{e}_{(\mu_1)} \otimes \cdots \otimes \mathbf{e}_{(\mu_k)} \otimes \varepsilon^{(\nu_1)} \otimes \cdots \otimes \varepsilon^{(\nu_\ell)}$.

Since the basis vectors are composed of all zeroes except for a single "1", all of the information is contained in the components $T^{\mu_1 \cdots \mu_\ell}_{\nu_1 \cdots \nu_k}$. Thus, we loosely say that $\mathcal{T}$ is isomorphic to the vector space of (n^k x n^ℓ) matrices $T^{\mu_1 \cdots \mu_\ell}_{\nu_1 \cdots \nu_k}$.

As with (1, 1) tensors, we also have the formula

$$T\left(\omega^{(1)},\cdots,\omega^{(k)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)}\right)=T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\,\omega_{\mu_1}\cdots\omega_{\mu_k}\,W^{\nu_1}\cdots W^{\nu_\ell} \tag{14}$$

and its matrix form

$$T\left(\omega^{(1)},\cdots,\omega^{(k)},\boldsymbol{W}_{(1)},\cdots,\boldsymbol{W}_{(\ell)}\right)=\omega^{(1)}\otimes\cdots\otimes\omega^{(k)}\,\boldsymbol{T}\,\boldsymbol{W}_{(1)}\otimes\cdots\otimes\boldsymbol{W}_{(\ell)}\,. \tag{15}$$

where $\boldsymbol{T}$ is a $n^k x n^\ell$ matrix .

Also, $\boldsymbol{T}$ can be expressed as the tensor product of k vectors $\boldsymbol{T}_{(i)}$ and ℓ covectors $\boldsymbol{T}^{(j)}$:

$$\boldsymbol{T}=\boldsymbol{T}_{(k)}\otimes\cdots\otimes\boldsymbol{T}_{(1)}\otimes\boldsymbol{T}^{(\ell)}\otimes\cdots\otimes\boldsymbol{T}^{(1)}\,. \tag{16}$$

This makes sense because $\omega=\omega^{(1)}\otimes\cdots\otimes\omega^{(k)}$ is a length n^k row vector, $\boldsymbol{T}$ is a $n^k x n^\ell$ matrix, and $\boldsymbol{W}=\boldsymbol{W}_{(1)}\otimes\cdots\otimes\boldsymbol{W}_{(\ell)}$ is length n^ℓ column vector. So it is legitimate to perform the matrix operations $\omega\boldsymbol{T}\boldsymbol{W}$ even though matrices ω, $\boldsymbol{T}$, and $\boldsymbol{W}$ are themselves tensor products. The indices on $\boldsymbol{T}$ in (16) are in reverse numerical order because $\omega^{(k)}\boldsymbol{T}_{(k)}$ is performed 1st, reducing to a scalar, then $\omega^{(k-1)}\boldsymbol{T}_{(k-1)}$ reduces to a scalar, etc., and similarly $\boldsymbol{T}^{(1)}\boldsymbol{W}_{(1)}$ becomes a scalar, then $\boldsymbol{T}^{(2)}\boldsymbol{W}_{(2)}$, …, and last, $\boldsymbol{T}^{(\ell)}\boldsymbol{W}_{(\ell)}$. The various matrix and tensor products can be performed in any order.

Just because a tensor T has been defined on k covectors and ℓ vectors does not mean it cannot act on a subset of covectors and/or vectors. For example,

$T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\,\omega_{\mu_2}\cdots\omega_{\mu_k}\,W^{\nu_2}\cdots W^{\nu_\ell}=\left(scalar\right)T^{\mu_1}_{\nu_1}$ is a (1, 1) tensor.

Moreover, not only can a tensor act on covectors and vectors, but any tensor can act on any other tensor to produce a tensor as follows.

Definition. If $\boldsymbol{R} \overset{(12)}{=} R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell}\,\mathbf{e}_{(\mu_1)}\otimes\cdots\otimes\mathbf{e}_{(\mu_k)}\otimes\varepsilon^{(\nu_1)}\otimes\cdots\otimes\varepsilon^{(\nu_\ell)}$ is a (k, ℓ) tensor and $\boldsymbol{S} \overset{(12)}{=} S^{\mu_{k+1} \cdots \mu_{k+m}}_{\nu_{\ell+1} \cdots \nu_{\ell+n}}\,\mathbf{e}_{(\mu_{k+1})}\otimes\cdots\otimes\mathbf{e}_{(\mu_{k+m})}\otimes\varepsilon^{(\nu_{\ell+1})}\otimes\cdots\otimes\varepsilon^{(\nu_{\ell+n})}$ is an (m, n) tensor, define a **(k+m, ℓ+n) tensor** $\boldsymbol{T}=\boldsymbol{R}\otimes\boldsymbol{S}$ as

$$
\begin{aligned}
\boldsymbol{T} = \boldsymbol{R} \otimes \boldsymbol{S} &= R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} \, \boldsymbol{e}_{(\mu_1)} \otimes \cdots \otimes \boldsymbol{e}_{(\mu_k)} \otimes \boldsymbol{\varepsilon}^{(\nu_1)} \otimes \cdots \otimes \boldsymbol{\varepsilon}^{(\nu_\ell)} \\
&\qquad \otimes \; S^{\mu_{k+1} \cdots \mu_{k+m}}_{\nu_{\ell+1} \cdots \nu_{\ell+n}} \, \boldsymbol{e}_{(\mu_{k+1})} \otimes \cdots \otimes \boldsymbol{e}_{(\mu_{k+m})} \otimes \boldsymbol{\varepsilon}^{(\nu_{\ell+1})} \otimes \cdots \otimes \boldsymbol{\varepsilon}^{(\nu_{\ell+n})} \\
&\equiv R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} \; S^{\mu_{k+1} \cdots \mu_{k+m}}_{\nu_{\ell+1} \cdots \nu_{\ell+n}} \boldsymbol{e}_{(\mu_1)} \otimes \cdots \otimes \boldsymbol{e}_{(\mu_{k+m})} \otimes \boldsymbol{\varepsilon}^{(\nu_1)} \otimes \cdots \otimes \boldsymbol{\varepsilon}^{(\nu_{\ell+n})} \\
&= T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} \boldsymbol{e}_{(\mu_1)} \otimes \cdots \otimes \boldsymbol{e}_{(\mu_{k+m})} \otimes \boldsymbol{\varepsilon}^{(\nu_1)} \otimes \cdots \otimes \boldsymbol{\varepsilon}^{(\nu_{\ell+n})}
\end{aligned}
\tag{17}
$$

where $T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} = R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} S^{\mu_{k+1} \cdots \mu_{k+m}}_{\nu_{\ell+1} \cdots \nu_{\ell+n}}$.

Then,

$$
\begin{aligned}
&\boldsymbol{R} \otimes \boldsymbol{S}\left(\boldsymbol{\omega}^{(1)}, \cdots, \boldsymbol{\omega}^{(k+\ell)}, \boldsymbol{W}_{(1)}, \cdots, \boldsymbol{W}_{(m+n)}\right) = \boldsymbol{T}\left(\boldsymbol{\omega}^{(1)}, \cdots, \boldsymbol{\omega}^{(k+\ell)}, \boldsymbol{W}_{(1)}, \cdots, \boldsymbol{W}_{(m+n)}\right) \\
&\overset{(14)}{=} \; T^{\mu_1, \cdots, \mu_{k+\ell}}_{\nu_1, \cdots, \nu_{m+n}} \omega_{\mu_1} \cdots \omega_{\mu_{k+\ell}} W^{\nu_1} \cdots W^{\nu_{m+n}} = R^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_\ell} S^{\mu_{k+1} \cdots \mu_{k+m}}_{\nu_{\ell+1} \cdots \nu_{\ell+n}} \omega_{\mu_1} \cdots \omega_{\mu_{k+\ell}} W^{\nu_1} \cdots W^{\nu_{m+n}}.
\end{aligned}
$$

This can also be viewed as tensor ***R*** acting on tensor ***S*** to produce tensor ***T*,** or as tensor ***S*** acting on tensor ***R*** to produce tensor ***T*.**

9 Examples

Example 1. A (2,2) tensor is a quad-linear map

$$T:\mathcal{V}^*\times\mathcal{V}^*\times\mathcal{V}\times\mathcal{V}\to\mathbf{F}:T(\omega,\xi,\boldsymbol{V},\boldsymbol{W})\overset{(5)}{=}\omega_{\mu_1}\xi_{\mu_2}T^{\mu_1\mu_2}_{\ \ \ \ \nu_1\nu_2}V^{\nu_1}W^{\nu_2}.$$

To neutralize covectors ω and ξ, they must be multiplied on their right by two vectors. To neutralize vectors $\boldsymbol{V}$ and $\boldsymbol{W}$, they must be multiplied on their left by two row vectors. Thus, a (2,2) tensor $\boldsymbol{T}$ is the tensor product of the tensor product of two vectors with the tensor product of two covectors:

$$T(\omega,\xi,\boldsymbol{V},\boldsymbol{W})=\omega\otimes\xi\,\boldsymbol{T}\,\boldsymbol{V}\otimes\boldsymbol{W}$$

where

$$\boldsymbol{T}=\boldsymbol{T}_{(\xi)}\otimes\boldsymbol{T}_{(\omega)}\otimes\boldsymbol{T}^{(W)}\otimes\boldsymbol{T}^{(V)}.$$

Example 2. Quantum Mechanics. $\mathbf{F}=\mathbb{C}$.
Alice and Bob each measure the spin state of an electron. Let $\mathrm{S_A}$ represent Alice's state space, the vector space of all her states. Let S_B represent Bob's state space. Let $|\boldsymbol{A}\rangle$ and $|\boldsymbol{B}\rangle$ be state vectors in $\mathrm{S_A}$ and $\mathrm{S_B}$, respectively, and let $\{\,|\boldsymbol{a}\rangle\,\}$ and $\{\,|\boldsymbol{b}\rangle\,\}$ be bases for $\mathrm{S_A}$ and $\mathrm{S_B}$. Then $\mathrm{S_{AB}}=\mathrm{S_A}\otimes\mathrm{S_B}$ is the vector space whose basis is the set of tensor product states $|\boldsymbol{ab}\rangle\equiv|\boldsymbol{a}\rangle\otimes|\boldsymbol{b}\rangle$. These basis objects are vectors since they are tensor products of vectors. Similarly,

if $|\boldsymbol{A}\rangle=\sum_a\alpha_a|a\rangle$ and $|\boldsymbol{B}\rangle=\sum_b\beta_b|\boldsymbol{b}\rangle$, then
$|\boldsymbol{AB}\rangle=|\boldsymbol{A}\rangle\otimes|\boldsymbol{B}\rangle=\sum_{\boldsymbol{a}}\sum_{\boldsymbol{b}}\alpha_{\boldsymbol{a}}\beta_{\boldsymbol{b}}|\boldsymbol{ab}\rangle\in\mathrm{S_{AB}}$ are vectors.
However, $\mathrm{S_{AB}}$ is larger than the set of vectors $\{\,|\boldsymbol{AB}\rangle\,\}$. The singlet state is a vector of the form $|\boldsymbol{\Psi}\rangle=\sum_{\boldsymbol{a}}\sum_{\boldsymbol{b}}\psi_{\boldsymbol{ab}}|\boldsymbol{ab}\rangle\in\mathrm{S_{AB}}$ that cannot be expressed as $|\boldsymbol{A}\rangle\otimes|\boldsymbol{B}\rangle$. We give a more detailed example of this behavior next.

Example 3. General Relativity. $\mathbf{F}=\mathbb{R}$.

Let $\boldsymbol{e}_{(0)}=\begin{pmatrix}i\\0\\0\\0\end{pmatrix}$, $\boldsymbol{e}_{(1)}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$, $\boldsymbol{e}_{(2)}=\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$, and $\boldsymbol{e}_{(3)}=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$ be a basis for spacetime, S. Let $\varepsilon^{(\nu)}=\boldsymbol{e}_{(\nu)}^{\mathrm{T}}$. For example, $\varepsilon^{(0)}=\begin{pmatrix}i&0&0&0\end{pmatrix}$.

Then $\left\{ \mathbf{e}_{(\mu)} \otimes \varepsilon^{(\nu)} \right\}_{\mu,\nu=0}^{3}$ is a basis for $S^2 = S \otimes S$, where elements of the basis are of the form

$$\mathbf{e}_{(\mu)}^{(\nu)} \equiv \mathbf{e}_{(\mu)} \otimes \boldsymbol{\varepsilon}^{(\nu)} = \begin{pmatrix} 0 & \vdots & 0 \\ \cdots & e^{\mu}\varepsilon_{\nu} & \cdots \\ 0 & \vdots & 0 \end{pmatrix}.$$

Define the metric tensor

$$\boldsymbol{\eta} = \left(\eta_{\nu}^{\mu}\right) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$\boldsymbol{\eta} = \sum_{\mu}\sum_{\nu} \eta_{\nu}^{\mu}\, \mathbf{e}_{(\mu)} \otimes \varepsilon^{(\nu)}$ is a member of S^2 that cannot be expressed a product tensor. That is, let $\boldsymbol{\alpha} = \sum_{\mu} \alpha^{\mu}\, \mathbf{e}_{(\mu)}$ and $\boldsymbol{\beta} = \sum_{\nu} \beta_{\nu} \boldsymbol{\varepsilon}^{(\nu)}$. Then $\boldsymbol{\eta} \neq \boldsymbol{\alpha} \otimes \boldsymbol{\beta}$:

Suppose

$$\boldsymbol{\eta} = \boldsymbol{\alpha} \otimes \boldsymbol{\beta} = \sum_{\mu}\sum_{\nu} \alpha^{\mu}\beta_{\nu}\, \mathbf{e}_{(\mu)} \otimes \boldsymbol{\varepsilon}^{(\nu)} = \begin{pmatrix} \alpha^0\beta_0 & \cdots & \alpha^0\beta_3 \\ \vdots & & \vdots \\ \alpha^3\beta_0 & \cdots & \alpha^3\beta_3 \end{pmatrix}.$$

Then

$$-1 = \alpha^0\beta_0 \quad \Rightarrow \quad \alpha^0 \neq 0 \quad \text{and} \quad 1 = \alpha^1\beta_1 \quad \Rightarrow \quad \beta_1 \neq 0.$$

Yet,

$$\alpha^0\beta_1 = 0,$$

a contradiction.

So the space S^2 generated by the basis $\left\{ \mathbf{e}_{(\mu)} \otimes \varepsilon^{(\nu)} \right\}_{\mu,\nu=0}^{3}$ is larger than the space of product vectors $\left\{ \boldsymbol{\alpha} \otimes \boldsymbol{\beta} \right\}$. Using quantum mechanics terminology we would say that the metric tensor is **entangled.**

Also, as mentioned, some books may label the metric tensor as the matrix $\eta_{\mu\nu}$. I prefer to write $\eta_{\mu\nu}$ as a (1 x 16) covector :

$$\begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

I would express $\eta^{\mu\nu}$ as a (16 x 1) column vector, and then the product $\eta_{\mu\nu}\,\eta^{\mu\nu}$ would be a scalar, not another matrix as would be the case if both tensors were matrices.

Finally, the metric tensor enables compact notation for things like the space-time interval between two events:

$$s^2 = -\left(c\Delta t\right)^2 + \left(\Delta x\right)^2 + \left(\Delta y\right)^2 + \left(\Delta z\right)^2.$$

To write this, we let $\boldsymbol{\Delta x} \equiv \begin{pmatrix} c\,\Delta t \\ \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}$. Then $s^2 = \left(\boldsymbol{\Delta x}\right)^{\mathsf{T}} \boldsymbol{\eta} \left(\boldsymbol{\Delta x}\right) = \eta^{\mu}_{\nu} \left(\Delta x\right)^{\mathsf{T}}_{\mu} \left(\Delta x\right)^{\nu}$.