Gravity 1

Asaf Arzi

November 12, 2022

Mathematical Supplement - Tensors

Abstract

Tensors are fundamental and practical objects in many areas of physics. In any field theory and especially in general relativity, matter, geometry and symmetries are formalized by tensor fields. This paper present the basic mathematical definitions and operations of tensors which are relevant for physics students. "Basic" here means that this paper treat tensors at the pure algebraic level, as a subject of linear algebra. Dealing with tensor **fields** would be an additional step, since they require also a differential structure. Also, we deal mainly with a pure vector space with no inner product which is also one step further, so the metric tensor is treated only in the appendix.

Instead, we want to clarify the equivalence of two different approaches to **define** tensors. One is the "math" approach, as a multilinear map[1]. It is presented in unit 2. This approach has the advantages of a straightforward definition and basis-free, "geometrical", formulation and conception. The second approach is the "physics" approach, as an array (written as a symbol with indices) obeying some transformation law. It is presented in unit 3. This approach is most important for calculations and relation to symmetries. Although "less abstract"[2] in practice, conceptually it is actually much harder to digest when presented first. Unit 1 prepares the ground for both approaches. This paper wishes to help us understand the expressions and operations we make in tensor calculations.

[1]There are also more complicated and abstract definitions in abstract algebra that are not useful to us.

[2]This approach also has a more abstract and geometrical formulation in the field of bundle theory.

Contents

1 Vectors and Covectors

1.1 The Dual Space

Let V be a finite dimensional vector space over the real numbers R. Its *dual space*, denoted V^*, consists of all **linear** maps from V to R (considered as a vector space)

$$V^* := \{\omega : V \to R\}$$

such that

$$\omega(v + u) = \omega(v) + \omega(u) \tag{1}$$

and

$$\omega(\lambda v) = \lambda\omega(v) \tag{2}$$

for all $v, u \in V$, $\lambda \in R$.

We can inherent from R an addition and scalar multiplication to the dual space

$$(\alpha + \beta)(v) := \alpha(v) + \beta(v) \tag{3}$$

$$(\lambda\alpha)(v) := \lambda\alpha(v) \tag{4}$$

Where $\alpha, \beta \in V^*$, $v \in V$, $\lambda \in R$.

One can show that with these, the dual space also constitutes a vector space structure. Hence, an element of V^* is called a *covector*[3].

Suppose we chose a basis for V, $\{e_1, e_2, ...e_{dimV}\}$. Note - we label the basis vectors of V with an index **downstairs**. Still, each e_i is a **vector**. It is convenient to choose a basis for V^* $\{\theta^1, \theta^2, ..., \theta^{dimV}\}$[4] such that

$$\theta^i(e_j) = \delta^i_j \tag{5}$$

[5]Such a basis is called the *dual basis* of the dual space. Note - we label the basis vectors of V^* with an index **upstairs**. Still, each θ^i is a **covector**.

Lets see why it is a nice choice, and how everything is tied together.

We write a vector v as a linear combination of the basis $\{e_i\}$

$$v = \sum_{i=1}^{dimV} v^i e_i \tag{6}$$

Note - the $\{v^i\}$ are the **components** of the vector v, they are real numbers, and we put their index **upstairs**.

We write a covector ω as a linear combination of the basis $\{\theta^i\}$

$$\omega = \sum_{i=1}^{dimV} \omega_i\theta^i \tag{7}$$

---

[3]It is also called a dual vector, or a (linear) functional.

[4]$dimV^* = dimV$

[5]For the Kronecker delta symbol the horizontal placing of indices does not matter $\delta^i_j \equiv {\delta^i}_j = {\delta_j}^i$

Note - the $\{\omega_i\}$ are the **components** of the covector ω, they are real numbers, and we put their index **downstairs**.

Recall that the dual vector is a linear map - its input is a vector and the output is a scalar. Let us evaluate $\omega(v)$.

$$\begin{aligned}\omega(v) &= \left(\sum_{i=1}^{dimV} \omega_i \theta^i\right)(v) = \sum_{i=1}^{dimV} \omega_i \theta^i(v) = \sum_{i=1}^{dimV} \omega_i \theta^i(\sum_{j=1}^{dimV} v^j e_j) \\ &= \sum_{i=1}^{dimV}\sum_{j=1}^{dimV} \omega_i v^j \theta^i(e_j) = \sum_{i=1}^{dimV}\sum_{j=1}^{dimV} \omega_i v^j \delta^i_j = \sum_{i=1}^{dimV} \omega_i v^i \end{aligned} \tag{8}$$

where first we plugged in the covector expansion (7), then used the addition and scalar multiplication on V^*(3)(4), then plugged in the vector expansion (6), going to the second line by the linearity of the covectors as maps (1)(2), and then using the dual basis definition (5).

Everything will look much cleaner if we adopt the *Einstein summation convention*, that is - any two repeated indices are summed over. It means that we will not write the sum signs anymore. Notice two things: First, that the indices conventions are such that the repeated indices are always one up and one down; Second, that we can make the sums transparent without harm only because of the linearity of the maps.

We conclude that

$$\omega(v) = \omega_i v^i \tag{9}$$

Covector acting on a vector

In a linear algebra course we represent vector components as a "column vector" $\begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix}$ and covector components as a "row vector" $\begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \end{pmatrix}$ and calculate $\omega(v)$ by inventing matrix multiplication algorithm

$$\begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \omega_1 v^1 + \omega_2 v^2 + \omega_3 v^3. \tag{10}$$

We see that this is identical to (9). It is just how a covector acts on a vector linearly, using any basis for V and its dual basis for V^*. Remark: This is NOT an inner product between two vectors, which does not even exists on a pure vector space. Also, (10) is sometimes written as $\omega^T v$, as if ω is secretly a vector who got transposed, but a covector is not a "transpose" of some vector, it was born as a covector.

Physics example: in quantum mechanics we denote the vectors as ket and covectors as bra, and the application of a covector on a vector is the bracket.

1.2 Change of Basis

Question: If we expand a vector v in some basis $\{e_i\}$, and then change to another basis $\{\widetilde{e}_i\}$, how does the vector components v^i transform? Let us have a look.

We change the old basis $\{e_i\}$ to a new basis $\{\widetilde{e}_i\}$ by a change-of-basis matrix J

$$\widetilde{e}_i = J^j{}_i e_j \tag{11}$$

The only requirment on J is that it is invertible, so $J^i{}_k \left(J^{-1}\right)^k{}_j = \delta^i_j$. The old basis in term of the new one is

$$e_i = \left(J^{-1}\right)^j{}_i \widetilde{e}_j \tag{12}$$

We can write a vector in the two different bases and compare

$$v = \widetilde{v}^i \widetilde{e}_i = v^j e_j = v^j \left(J^{-1}\right)^i{}_j \widetilde{e}_i \tag{13}$$

where we plugged in (12). Notice that when summing over an index we are free to choose how to name it (it is a *dummy index*). We used it here so we can see clearly that the **vector components** transformation is

$$\boxed{\widetilde{v}^i = \left(J^{-1}\right)^i{}_j v^j} \tag{14}$$

Vector components transformation

(14) means that the vector components transform under change of basis the opposite way that the basis transform (with the inverse matrix). The basis transformation (11) is taken to be the agreed reference, so we say that the basis vectors are *covariant* (transform like themselves) and the vector components are *contravariant* (transform the opposite way). This is of course nothing new, it is the statement that the abstract vector is the same in any basis. Let us make a consistency check while using the previous equations and index notation:

$$v = \widetilde{v}^i \widetilde{e}_i = v^j \left(J^{-1}\right)^i{}_j J^k{}_i e_k = v^j \delta^k_j e_k = v^j e_j = v \tag{15}$$

Remark: the indexed components are just commuting real numbers, so their order does not matter, for example $v^j J^i{}_j = J^i{}_j v^j$.

Now let us find how the dual basis of V^* transform when we change the basis of V. We want to find the matrix transformation M

$$\widetilde{\theta}^i = M^i{}_j \theta^j \tag{16}$$

The requierment for $\{\widetilde{\theta}^i\}$ is that this new basis is dual to the new $\{\widetilde{e}_i\}$, so that (5) still holds. Performing the transformations (11)(16) we find

$$\delta^i_j = \widetilde{\theta}^i(\widetilde{e}_j) = M^i{}_l \theta^l (J^k{}_j e_k) = M^i{}_l J^k{}_j \theta^l(e_k) = M^i{}_l J^k{}_j \delta^l_k = M^i{}_k J^k{}_j \tag{17}$$

Therefore $M = J^{-1}$. We conclude

$$\widetilde{\theta}^i = \left(J^{-1}\right)^i{}_j \theta^j \tag{18}$$

Exercise: Show that the **covector components** transform as

$$\boxed{\widetilde{\omega}_i = J^j{}_i \omega_j} \tag{19}$$

Covector components transformation

They transform opposite to the dual basis, for the same reason the vector components transform opposite to the original basis vectors.

The summary of these rules is as follows: An object with an index **down** is **covariant** (transforms with J) and an object with an index **up** is **contravariant** (transforms with J^{-1}). Hence, the basis vectors e_i and the covector components ω_i are covariant while the dual basis θ^i and the vector components v^i are contravariant. It means that when we sum over all up and down indices the result does not depend on the choice of basis (it is invariant) since the transformations cancel. This is the power of the up-down index notation. Check for example that (9) is indeed invariant. Caveat: This rule holds only **if** the "indexed objects" are indeed some kind of tensors (which will be defined later).

Picture of vector/covector and physics We now see that the difference between vectors and covectors is how they transform. Lets think of the most simple physical example, in one dimension. Say I have a rod and I want to measure its length. I have 1-meter sticks, and I find that exactly two of them together match the rod. So the rod has length of 2 meters. The 1-meter stick is my basis vector, and the number 2 is the vector component. Now I perform the measurement with 1-cm sticks. Now of course I need to combine 200 of the new sticks to match the rod's length. So I say that I changed my basis vector by factor of 1/100 and thus the vector component transformed by a factor of 100, just the opposite. The rod has the same length regardless how I measure it.

From this simple example we conclude that a quantity with units of length is a vector. On the contrary, a quantity with units of $(length)^{-1}$ or $(time)^{-1}$ is by the same analysis a covector ($1\frac{1}{sec} = 60\frac{1}{min}$) so a frequency or a line density transform the same as the basis. Geometrically, for example in three dimensions, we can represent a vector by an arrow and a covector as a stack of parallel planes[6]. The direction of the covector is perpendicular to the planes and its magnitude is their line-density. Acting with a covector on a vector is putting the arrow on top of the planes and the result is the number of planes intersecting the arrow[7]. Maybe the most famous covector is the wave "vector" k - the "spatial frequency" of a plane wave. It is the stack of planes of constant phase. Its action on a displacement vector x give the total phase accumulated by the front over a distance x, $\phi = k(x) = k_i x^i$.

[6]In two dimensions these are lines, and in general d dimensions these are $d-1$ hyperplanes.
[7]Not only integer, can be any real number.

Tensors

2 Tensors as Multi-Linear Maps

2.1 First Tensor Definition

Definition: A $\begin{pmatrix} p \\ q \end{pmatrix}$-tensor T is a **multilinear** map

$$T : \underbrace{V^* \times ... \times V^*}_{p} \times \underbrace{V \times ... \times V}_{q} \to R \tag{20}$$

It means that it receives p covectors and q vectors and return a scalar, so $T(\omega_1, ..., \omega_p, v_1, ..., v_q) \in R$.[8] Multilinearity means it is linear in each argument. For example, a bi-linear map $T : V^* \times V \to R$ satisfies

$$T(\lambda_1\alpha + \lambda_2\beta, \lambda_3 v + \lambda_4 u) = \lambda_1\lambda_3 T(\alpha, v) + \lambda_1\lambda_4 T(\alpha, u) + \lambda_2\lambda_3 T(\beta, v) + \lambda_2\lambda_4 T(\beta, u) \tag{21}$$

for all $\alpha, \beta \in V^*$ $v, u \in V$ and $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in R$.

$\begin{pmatrix} p \\ q \end{pmatrix}$ is called the order/rank/degree/type of the tensor, and sometimes people call $p+q$ the order/rank/degree of the tensor. As we did for covectors, by inheriting the addition and scalar multiplication from R, we can add and scale point-wise $\begin{pmatrix} p \\ q \end{pmatrix}$ tensors. The set of all tensors of a specific type $\begin{pmatrix} p \\ q \end{pmatrix}$ constitutes a vector space, denoted $\mathcal{T}^p_q$ with $dim\mathcal{T}^p_q = (dimV)^{p+q}$.

Special types

Total rank 0:

- $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$-tensor is a **scalar**, a real number, and $R = \mathcal{T}^0_0$.

Total rank 1:

- A $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$-tensor is a linear map $T : V \to R$. Thus, a **covector** is a special type of tensor, namely a $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$-tensor, and $V^* = \mathcal{T}^0_1$.

- For finite dimensional vector spaces $(V^*)^* = V$, that is, the vector space is the dual of the dual space. Thus a vector is a linear map $v : V^* \to R$, so a **vector** is a $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$-tensor, and $V = \mathcal{T}^1_0$.

[8] The indices here are just names for the different covectors and vectors, not components.

Total rank 2:

- A $\binom{1}{1}$-tensor is a multilinear map $T : V^* \times V \to R$, so $T(\omega, v) \in R$. It is equivalent to say that $T(\ , v)$ is waiting to eat a covector to produce a scalar, so $T(\ , v) : V^* \to R$. Therefore $T(\ , v)$ is a vector. It means that feeding T with a vector v produces linearly another vector $T(\ , v)$, hence $T : V \to V$. The conclusion is that a $\binom{1}{1}$-tensor is a **linear operator**[9] on V. We know it from linear algebra and can represent it as a square matrix. In the same way we can say that $\binom{1}{1}$-tensor is a linear map $T : V^* \to V^*$, a linear operator on V^*. To summarize, this type of tensor maps a vector to a vector, or a covector to a covector, or a vector and covector to a scalar. In matrix form

$$\overset{T}{\begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}} \overset{v}{\begin{pmatrix} \\ \\ \\ \end{pmatrix}} = \overset{T(\ ,v)}{\begin{pmatrix} \\ \\ \\ \end{pmatrix}}$$

,

$$\overset{\omega}{\begin{pmatrix} & & \end{pmatrix}} \overset{T}{\begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}} = \overset{T(\omega,\ )}{\begin{pmatrix} & & \end{pmatrix}}$$

,

$$\overset{\omega}{\begin{pmatrix} & & \end{pmatrix}} \overset{T}{\begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}} \overset{v}{\begin{pmatrix} \\ \\ \\ \end{pmatrix}} = \overset{T(\omega,v)}{()}$$

- A $\binom{0}{2}$-tensor is a multilinear map $T : V \times V \to R$, so $T(v, u) \epsilon R$. This is called in linear algebra a **bilinear form**. In this case $(T(v,\ ) : V \to R) \in V^*$. So in contrast to the linear operator above, this type of tensor eats a vector and returns a covector. It is a linear map $T : V \to V^*$. Therefore it is dangerous to think of it as a matrix in the usual sense. An important special case is the inner product of two vectors, which in this tensor language is called a **metric tensor**.
- A $\binom{2}{0}$-tensor is a multilinear map $T : V^* \times V^* \to R$, so $T(\alpha, \beta) \in R$. Equivalently, it is a linear map $T : V^* \to V$.

2.2 Tensor Components

Now we are ready to write the components of abstract tensors . For simplicity of notation let us start with second rank tensors.

[9] Also called an endomorphism , and $\mathcal{T}^1_1 = End(V)$.

- A $\binom{0}{2}$-tensor T action on two vectors v, u is

$$T(v,u) = T(v^i e_i, u^j e_j) = v^i u^j T(e_i, e_j) \equiv v^i u^j T_{ij} \tag{22}$$

where

$$T_{ij} := T(e_i, e_j) \tag{23}$$

are the $\binom{0}{2}$-tensor T components with respect to the $\{e_i\}$ basis.

- A $\binom{2}{0}$-tensor T action on two covectors α, β is

$$T(\alpha,\beta) = T(\alpha_i \theta^i, \beta_j \theta^j) = \alpha_i \beta_j T(\theta^i, \theta^j) \equiv \alpha_i \beta_j T^{ij} \tag{24}$$

where

$$T^{ij} := T(\theta^i, \theta^j) \tag{25}$$

are the $\binom{2}{0}$-tensor T components with respect to the $\{\theta^i\}$ basis.

- A $\binom{1}{1}$-tensor action T on a covector ω and a vector v is

$$T(\omega, v) = T(\omega_i \theta^i, v^j e_j) = \omega_i v^j T(\theta^i, e_j) \equiv \omega_i v^j T^i{}_j \tag{26}$$

where

$$T^i{}_j := T(\theta^i, e_j) \tag{27}$$

are the $\binom{1}{1}$-tensor T components with respect to the $\{\theta^i\}$ and $\{e_i\}$ basis.

In general we apply a $\binom{p}{q}$-tensor on p covectors $\{\omega_i\}_{i=1}^{i=p}$ and q vectors $\{v_j\}_{j=1}^{j=q}$ and expand them (7) (6).

$$\begin{aligned} T(\omega_1, ..., \omega_p, v_1, ..., v_q) &= T\left((\omega_1)_{i_1} \theta^{i_1}, ..., (\omega_p)_{i_p} \theta^{i_p}, (v_1)^{j_1} e_{j_1}, ..., (v_q)^{j_q} e_{j_q}\right) \\ &= (\omega_1)_{i_1} ... (\omega_p)_{i_p} (v_1)^{j_1} ... (v_q)^{j_q}\, T\left(\theta^{i_1}, ..., \theta^{i_p}, e_{j_1}, ..., e_{j_q}\right) \\ &= (\omega_1)_{i_1} ... (\omega_p)_{i_p} (v_1)^{j_1} ... (v_q)^{j_q}\, T^{i_1 \dots i_p}{}_{j_1 \dots j_q} \end{aligned} \tag{28}$$

where

$$T^{i_1 \dots i_p}{}_{j_1 \dots j_q} := T\left(\theta^{i_1}, ..., \theta^{i_p}, e_{j_1}, ..., e_{j_q}\right) \tag{29}$$

Tensor components

are the $\binom{p}{q}$-tensor components with respect to the $\{\theta^i\}$ and $\{e_i\}$ basis.

Summary:

1. The components of a tensor are extracted by applying it to all the basis vectors and dual basis covectors it needs in order to produce a real number, as in (29). Notice the matching of the indices there, p up and q down.

2. The application of some p covectors and some q vectors to a $\binom{p}{q}$-tensor is written in components as $p+q$ summed indices between all of them and the tensor, as in (30).

$$T(\omega_1, ..., \omega_p, v_1, ..., v_q) = T^{i_1 \dots i_p}{}_{j_1 \dots j_q} \, (\omega_1)_{i_1} \dots (\omega_p)_{i_p} \, (v_1)^{j_1} \dots (v_q)^{j_q} \tag{30}$$

Question: What then does the expression $T^i_j v^j$ represent? Notice it has one *free index* up. Answer:

$$(T(\ ,v))^i = T(\theta^i, v) = T(\theta^i, v^j e_j) = v^j T(\theta^i, e_j) = T^i{}_j v^j \tag{31}$$

It is the i-th component of $T(\ ,v)$ - the vector produced by acting with the $\binom{1}{1}$-tensor T on the vector v. The free i index is up, as it should for vector components.

Action of a $\binom{0}{2}$-tensor T on a vector v in its first argument produces a covector $T(v,\ )$ with components

$$(T(v,\ ))_i = T(v, e_i) = T(v^j e_j, e_i) = v^j T(e_j, e_i) = T_{ji} v^j \tag{32}$$

while inserting v in its second argument yields

$$(T(\ ,v))_i = T(e_i, v) = T(e_i, v^j e_j) = v^j T(e_i, e_j) = T_{ij} v^j \tag{33}$$

The idea is to translate index notation to abstract notation and vise versa so that we know what is going on.

2.3 Tensor Product

We can make a new tensor from existing tensors, by introducing a *tensor product* $\otimes$.

For simplicity we shall introduce first a special case, namely a tensor product between two vectors, and then present the general tensor product of two tensors.

Tensor product of two vectors

Let $v, u \in V$. The tensor product $v \otimes u$ is a $\binom{2}{0}$-tensor defined by its action on two arbitrary covectors (α, β) as

$$(v \otimes u)\,(\alpha, \beta) := v\,(\alpha)\,u\,(\beta) \tag{34}$$

It is also defined such that a scalar λ is "blind" to the tensor product

$$\lambda\left(v\otimes u\right)=\left(\lambda v\right)\otimes u=v\otimes\left(\lambda u\right) \tag{35}$$

The tensor product also allows us to write a basis for each tensor space. In this case, given a basis $\{e_i\}$ for V, a basis for $\mathcal{T}_0^2$ is

$$e_i\otimes e_j \tag{36}$$

$\mathcal{T}_0^2$ is generated by this basis, so it is also called a *tensor product space* and denoted as $\mathcal{T}_0^2\equiv V\otimes V$. We can expand a general $\begin{pmatrix}2\\0\end{pmatrix}$-tensor T in this basis as

$$T=T^{ij}e_i\otimes e_j \tag{37}$$

where T^{ij} are the components of T. Lets see that these are the same components as defined in (29)

$$T\left(\theta^i,\theta^j\right)=\left(T^{kl}e_k\otimes e_l\right)\left(\theta^i,\theta^j\right)=T^{kl}e_k\left(\theta^i\right)e_l\left(\theta^j\right)=T^{kl}\delta_k^i\delta_j^j=T^{ij} \tag{38}$$

where we used the definitions of the tensor product (34)(35) and the dual basis (5) (and that $V=\left(V^*\right)^*$).

The components of the tensor product of two vectors is

$$\left(v\otimes u\right)^{ij}=\left(v\otimes u\right)\left(\theta^i,\theta^j\right)=v\left(\theta^i\right)u\left(\theta^j\right)=v^iu^j \tag{39}$$

Acting on two covectors

$$\begin{aligned}\left(v\otimes u\right)\left(\alpha,\beta\right)=&\left(\left(v^ie_i\right)\otimes\left(u^je_j\right)\right)\left(\alpha_k\theta^k,\beta_l\theta^l\right)\\=&v^iu^j\alpha_k\beta_le_i\left(\theta^k\right)e_j\left(\theta^l\right)=v^iu^j\alpha_k\beta_l\delta_i^k\delta_j^l=v^iu^j\alpha_i\beta_j\end{aligned} \tag{40}$$

Acting on one covector produces a new vector

$$\left(v\otimes u\right)\left(\alpha,\ \right)=v^ie_i\left(\alpha_k\theta^k\right)u=v^i\alpha_iu^je_j \tag{41}$$

Consider $dimV=2$. The components of $v\otimes u$ in matrix representation according to (39) are

$$\left(v\otimes u\right)^{ij}=\begin{pmatrix}v^1\\v^2\end{pmatrix}\begin{pmatrix}u^1&u^2\end{pmatrix}=\begin{pmatrix}v^1u^1&v^1u^2\\v^2u^1&v^2u^2\end{pmatrix} \tag{42}$$

From the $dimV=2$ basis vectors $e_1=\begin{pmatrix}1\\0\end{pmatrix}$ and $e_2=\begin{pmatrix}0\\1\end{pmatrix}$ we construct the $\left(dimV\right)^2=4$ basis tensors

$$\left(e_1\otimes e_1\right)^{ij}=\begin{pmatrix}1&0\\0&0\end{pmatrix}\quad\left(e_1\otimes e_2\right)^{ij}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$$

$$\left(e_2\otimes e_1\right)^{ij}=\begin{pmatrix}0&0\\1&0\end{pmatrix}\quad\left(e_2\otimes e_2\right)^{ij}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

Notice that not any $\binom{2}{0}$-tensor can be written as a single tensor product of two vectors. For example $e_1 \otimes e_1 + e_2 \otimes e_2$ (represented by the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$) cannot be. A tensor that can be written as a product, namely $T^{ij} = v^i u^j$, is called a *pure tensor.* In quantum mechanics the Hilbert space of a composite system is a tensor product space, and such states are called pure states. A state which is not pure is called an entangled state.

General tensor product

The tensor product between a $\binom{p}{q}$-tensor T and a $\binom{r}{s}$-tensor S yields a $\binom{p+r}{q+s}$-tensor $T \otimes S$

$$\otimes : \mathcal{T}^p_q \times \mathcal{T}^r_s \to \mathcal{T}^{p+r}_{q+s}$$

It is defined by its action on arbitrary $p+r$ covectors $(\omega_1, ..., \omega_{p+r})$ and $q+s$ vectors $(v_1, ..., v_{q+s})$ as

$$(T \otimes S)\left(\omega_1, ..., \omega_{p+r}, v_1, ..., v_{q+s}\right) := T\left(\omega_1, ...\omega_p, v_1, ...v_q\right) S\left(\omega_{p+1}, ..., \omega_{p+r}, v_{q+1}, ..., v_{q+s}\right) \tag{43}$$

The components of $T \otimes S$ are

Components of a tensor product

$$\begin{aligned}(T \otimes S)^{i_1 \ldots i_p k_1 \ldots k_r}{}_{j_1 \ldots j_q l_1 \ldots l_s} &= T\left(\theta^{i_1}, ...\theta^{i_p}, e_{j_1}, ...e_{j_q}\right) S\left(\theta^{k_1}, ...\theta^{k_r}, e_{l_1}, ...e_{l_s}\right) \\ &= T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} S^{k_1 \ldots k_r}{}_{l_1 \ldots l_s}\end{aligned} \tag{44}$$

Given a basis for $\{e_i\}$ for V, a basis for $\mathcal{T}^p_q$ is

$$e_{i_1} \otimes ... \otimes e_{i_p} \otimes \theta^{j_1} \otimes ... \otimes \theta^{j_q} \tag{45}$$

A general $\binom{p}{q}$-tensor T expansion is

$$T = T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} e_{i_1} \otimes ... \otimes e_{i_p} \otimes \theta^{j_1} \otimes ... \otimes \theta^{j_q} \tag{46}$$

The $\binom{p}{q}$ *tensor product space* is the vector space generated by this basis, so is denoted also as

$$\mathcal{T}^p_q \equiv \underbrace{V \otimes ... \otimes V}_{p} \otimes \underbrace{V^* \otimes ... \otimes V^*}_{q} = \left\{ T : \underbrace{V^* \times ... \times V^*}_{p} \times \underbrace{V \times ... \times V}_{q} \underset{multilinear}{\longrightarrow} R \right\} \tag{47}$$

Contraction

Contraction is an operation that makes a $\binom{p}{q}$-tensor to a $\binom{p-1}{q-1}$-tensor by summation over one covariant index and one contravariant index. It is done by letting one of the dual basis covectors eat one of the basis vectors. It is a generalization of the trace.

Examples:

$$T^i{}_j e_i \otimes \theta^j \longmapsto T^i{}_j \cancel{e_i} \otimes \theta^j (e_i) = T^i{}_j \delta^j_i = T^i{}_i \tag{48}$$

$$T^i{}_{jkl} e_i \otimes \theta^j \otimes \theta^k \otimes \theta^l \longmapsto T^i{}_{jkl} \cancel{e_i} \otimes \theta^j \otimes \theta^k (e_i) \otimes \theta^l = T^i{}_{jkl} \delta^k_i \theta^j \otimes \theta^l = T^i{}_{jil} \otimes \theta^j \otimes \theta^l \tag{49}$$

The contraction of the i_k and j_l indices of a $\binom{p}{q}$-tensor

$$\begin{aligned}
T \longmapsto\ & T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} e_{i_1} \otimes \ldots \otimes \cancel{e_{i_k}} \otimes \ldots \otimes e_{i_p} \otimes \theta^{j_1} \otimes \ldots \otimes \theta^{j_l} (e_{i_k}) \otimes \ldots \otimes \theta^{j_q} \\
= \delta^{j_l}_{i_k} & T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} e_{i_1} \otimes \ldots \otimes \cancel{e_{i_k}} \otimes \ldots \otimes e_{i_p} \otimes \theta^{j_1} \otimes \ldots \otimes \cancel{\theta^{j_l}} \otimes \ldots \otimes \theta^{j_q} \\
= & T^{i_1 \ldots i_k \ldots i_p}{}_{j_1 \ldots i_k \ldots j_q} e_{i_1} \otimes \ldots \otimes \cancel{e_{i_k}} \otimes \ldots \otimes e_{i_p} \otimes \theta^{j_1} \otimes \ldots \otimes \cancel{\theta^{j_l}} \otimes \ldots \otimes \theta^{j_q}
\end{aligned} \tag{50}$$

We can make contractions between tensor products, for example

$$(T \otimes S)^{ijk}{}_{lmn} = T^{ij} S^k{}_{lmn} \longmapsto T^{ij} S^k{}_{lij} \tag{51}$$

Now we are ready to make the step to an alternative definition of tensors, by their components and how they transform under a change of basis.

3 Tensors as Arrays Obeying a Transformation Law

3.1 Change of basis

How do tensor components transform under change of basis (11) and dual basis (18)? Let us denote the new basis objects by a tag on top of the indices instead of a tilde. The components $T^{i_1 \ldots i_p}_{j_1 \ldots j_q}$ defined as the multilinear mapping of the basis and dual basis vectors transform as

$$\begin{aligned}
T^{i'_1 \ldots i'_p}{}_{j'_1 \ldots j'_q} &:= T\left(\theta^{i'_1}, \ldots, \theta^{i'_p}, e_{j'_1}, \ldots, e_{j'_q}\right) = T\left(\left(J^{-1}\right)^{i'_1}{}_{i_1} \theta^{i_1}, \ldots, \left(J^{-1}\right)^{i'_p}{}_{i_p} \theta^{i_p}, J^{j_1}{}_{j'_1} e_{j_1}, \ldots, J^{j_q}{}_{j'_q} e_{j_q}\right) \\
&= \left(J^{-1}\right)^{i'_1}{}_{i_1} \ldots \left(J^{-1}\right)^{i'_p}{}_{i_p} J^{j_1}{}_{j'_1} \ldots J^{j_q}{}_{j'_q} T\left(\theta^{i_1}, \ldots, \theta^{i_p}, e_{j_1}, \ldots, e_{j_q}\right) \\
&= \left(J^{-1}\right)^{i'_1}{}_{i_1} \ldots \left(J^{-1}\right)^{i'_p}{}_{i_p} J^{j_1}{}_{j'_1} \ldots J^{j_q}{}_{j'_q} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}
\end{aligned} \tag{52}$$

This is the most important formula in this paper:

$$T^{i'_1 \dots i'_p}_{ j'_1 \dots j'_q} = \left(J^{-1}\right)^{i'_1}_{ i_1} \dots \left(J^{-1}\right)^{i'_p}_{ i_p} J^{j_1}_{ j'_1} \dots J^{j_q}_{ j'_q} T^{i_1 \dots i_p}_{ j_1 \dots j_q} \tag{53}$$

$\binom{p}{q}$-tensor components transformation

The multilinearity of the tensor mapping result in a simple transformation law: Each index down transforms as a covector component (covariant) and each index up transforms as a vector component (contravariant) - overall $p+q$ simultaneous independent transformations.

We can also see it from the definition of $T^{i_1 \dots i_p}_{ j_1 \dots j_q}$ as the components of T with respect to a tensor basis $e_{i_1} \otimes \dots \otimes e_{i_p} \otimes \theta^{j_1} \otimes \dots \otimes \theta^{j_q}$. We write the tensor T in two tensor bases and compare

$$\begin{aligned}
T &= \left[T^{i'_1 \dots i'_p}_{ j'_1 \dots j'_q}\right] e_{i'_1} \otimes \dots \otimes e_{i'_p} \otimes \theta^{j'_1} \otimes \dots \otimes \theta^{j'_q} = T^{i_1 \dots i_p}_{ j_1 \dots j_q} e_{i_1} \otimes \dots \otimes e_{i_p} \otimes \theta^{j_1} \otimes \dots \otimes \theta^{j_q} \\
&= T^{i_1 \dots i_p}_{ j_1 \dots j_q} \left(\left(J^{-1}\right)^{i'_1}_{ i_1} e_{i'_1}\right) \otimes \dots \otimes \left(\left(J^{-1}\right)^{i'_p}_{ i_p} e_{i'_p}\right) \otimes \left(J^{j_1}_{ j'_1} \theta^{j'_1}\right) \otimes \dots \otimes \left(J^{j_q}_{ j'_q} \theta^{j'_q}\right) \\
&= \left[\left(J^{-1}\right)^{i'_1}_{ i_1} \dots \left(J^{-1}\right)^{i'_p}_{ i_p} J^{j_1}_{ j'_1} \dots J^{j_q}_{ j'_q} T^{i_1 \dots i_p}_{ j_1 \dots j_q}\right] e_{i'_1} \otimes \dots \otimes e_{i'_p} \otimes \theta^{j'_1} \otimes \dots \otimes \theta^{j'_q}
\end{aligned} \tag{54}$$

Comparing the components in the square brackets yields (53). This is exactly how we found the vector components transformation (15), only here we transform the tensor basis. The change of basis vectors yields a change in the dual basis, and since their tensor products make the tensor basis, it changes accordingly. The tensor components transform opposite to the tensor basis, just as in any vector space. The key thing here is that the tensor product is distributive, associative, and commutes with scalars, which mimic the multilinearity of tensor maps used in (52).

<u>Transformations for specific types of tensors</u>

- A scalar is a $\binom{0}{0}$-tensor and it has one component which is invariant.
- A vector is a $\binom{1}{0}$-tensor and of course, as in (14):

$$v^{i'} = \left(J^{-1}\right)^{i'}_{ i} v^i \tag{55}$$

In matrix notation, we would write a matrix J^{-1} times a column vector v, $\widetilde{v} = J^{-1}v$. For each row of the matrix (first index) we multiply and sum all its columns (second index) with the vecor components.

- A covector is a $\binom{0}{1}$-tensor and of course, as in (19):

$$\omega_{i'} = J^i_{ i'} \omega_i \tag{56}$$

In matrix notation, we would write a row vector ω times a matrix J , $\widetilde{\omega} = \omega J$. For each column of the matrix (second index) we multiply and sum all its rows (first index) with the covecor components.

- A $\binom{1}{1}$-tensor:

$$T^{i'}_{\ j'} = \left(J^{-1}\right)^{i'}_{\ i} J^{j}_{\ j'} T^{i}_{\ j} \tag{57}$$

The (ij) components of matrix multiplication AB is $(AB)^{i}_{\ j} = A^{i}_{\ k} B^{k}_{\ j}$. So, in order to see clearly how to write this with matrix multiplication rule, we write in the following order $T^{i'}_{\ j'} = \left(J^{-1}\right)^{i'}_{\ i} T^{i}_{\ j} J^{j}_{\ j'}$. Indeed, a linear operator transforms in matrix notation as

$$\widetilde{T} = J^{-1} T J \tag{58}$$

- A $\binom{0}{2}$-tensor:

$$T_{i'j'} = J^{i}_{\ i'} J^{j}_{\ j'} T_{ij} \tag{59}$$

The summation over the i index in $J^{i}_{\ i'} T_{ij}$ is row-row summation, and not column-row summation as in matrix multiplication algorithm. A matrix transpose switches the rows and columns, i.e., it switches the order of indices. Therefore, $J^{i}_{\ i'} T_{ij}$ can be written in matrix form matrix as $J^{T} T$. Indeed, a bilinear form transforms in matrix notation as

$$\widetilde{T} = J^{T} T J \tag{60}$$

Notice that the determinant of a $\binom{1}{1}$-tensor is invariant since

$$det(J^{-1}TJ) = det(J^{-1})det(T)det(J) = det(J)^{-1}det(T)det(J) = det(T) \tag{61}$$

while the determinant of a $\binom{0}{2}$-tensor[10] is basis dependent since

$$det(J^{T}TJ) = det(J^{T})det(T)det(J) = det(J)det(T)det(J) = (detJ)^{2}det(T) \tag{62}$$

3.2 Second Tensor Definition

Up to this point the route we took to present tensors was top down, meaning from the abstract (maps) to the concrete (components). We defined tensors as an abstract ("basis free", "geometrical") objects which as such **do not** transform under change of basis, rather only their **components** does in order to accommodate for the basis change. This is also a standard viewpoint we have about abstract vectors.

However, we can take the opposite route - a bottom-up approach. We start by presenting some object with $p + q$ indices $T^{i_1 \ldots i_p}_{\quad\ j_1 \ldots j_q}$ according to some rule or a formula made from other existing objects. At this point these are **not necessarily** the components of some abstract tensor. All we have at this point is an **array of numbers**, with no structure. By structure we mean how it

[10] However it is defined...

transforms when we change a basis. So in order for this array to make sense we need to **check** how it transforms. For example, we can define a **symbol** like the Levi-Civita symbol ϵ_{ijk} which has the same values in any basis. If we invent an array $T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$ **and** it transforms as (53), then the $T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$ are components of some abstract $\begin{pmatrix} p \\ q \end{pmatrix}$-tensor T. Why? Because this transformation rule encodes completely the multilinearity of the abstract tensor as a map. In other words, giving an array and the way it transforms under a change of basis, supplies all the information about the abstract object. If the transformation law is (53) then it is components of a tensor which we can construct.

Some important points:

• Terminology -

Up to now we used the term "tensor" in order to refer to the abstract object T which does not transform, and we used the term "tensor components" to the array $T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$, defined by (29). However, in the bottom-up approach we present now, which is usually the starting point in physics textbooks, the term "tensor" refers to the array $T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$, with the **requirement** that it obeys (53). In this language, if $T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$ transforms as (53) we would say that "$T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$ is a tensor", and if it does not then we would say that "$T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$ is not a tensor". These two definitions and terminologies are both legitimate and are equivalent. People often say that "a tensor is something that transforms like a tensor", which is of course nonsense, unless you know what they really mean, which is the following second definition of a tensor.

Definition: A tensor of type $\begin{pmatrix} p \\ q \end{pmatrix}$ is an assignment of an array $T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$ to each basis of a vector space, such that if we apply a change of basis then the array obeys the transformation law (53).

• Zero tensor -

The linear nature of the transformation (53) yields a crucial advantage to the usage of tensors. If a tensor is zero in one basis, i.e., all its components are zero, then this transformation law says it is zero in any basis. This is just like with the zero vector we know. So we can write equations of tensors in one basis, and they hold (the equal sign remains correct) in any other basis. For example, the equation $T_{ijk} = M_{ijk}$ is equivalent to $T_{ijk} - M_{ijk} = 0$. So the tensor $F_{ijk} := T_{ijk} - M_{ijk}$ is the zero tensor, $F_{ijk} = 0$. We would like to write our physical laws in such a way. In general an array with a different transformation law can be all zero's in one basis and not all zero's in another. The prime example we will meet are the "connection coefficients" Γ^i_{jk}, which obey a somewhat different transformation law. It does not mean they are not components of some abstract well defined geometrical object, it just means that this object is not a tensor - not a multilinear map[11].

• Restriction of basis transformations-

A general change of basis is made by the action of the J matrix. The only requirement on that matrix is that it must be invertible. In group theory language we say that it is an element of the General Linear group $J \in GL(n)$, where $n = dimV$. There are circumstances where we may wish to restrict the allowed basis transformations. For example, in an inner product vector space we may want to allow only orthogonal basis transformations, which correspond to rotations and reflections (preserve the inner product), or special orthogonal transformations which correspond only to proper rotations. In these cases we restrict J to be an element of the subgroup $O(n)$ or $SO(n)$ respectively. Then, our definition of a tensor is modified such that the transformation law of tensors is (53) with $J \in SO(n)$ for example. In this case we would say "T_{ij} is an $SO(n)$-tensor". A $GL(n)$-tensor is also a tensor with respect to any subgroup of $GL(n)$, but an object could be a $O(n)$-tensor and not a general tensor. In special relativity we restrict to the restricted Lorentz transformations $SO^+(3,1)$ which relate different inertial frames, and talk about "Lorentz tensors".

3.3 Canonical Examples

Exercise: Prove that the Kronecker delta symbol is a $\binom{1}{1}$-tensor.

The Kronecker delta **symbol** δ^i_j is defined as the array

$$\delta^i_j = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases} \tag{63}$$

in any basis. Again, the delta symbol is defined to have the same numerical values independnt of the basis, so we need to check if in addition to it being invariant it is also a tensor. In the definition we use here tensors are in general not invariant, they transform as (53). In this case we need to **prove** that

$$\delta^{i'}_{j'} = \left(J^{-1}\right)^{i'}{}_{i} J^{j}{}_{j'} \delta^i_j \tag{64}$$

hold. Start from the r.h.s

$$\left(J^{-1}\right)^{i'}{}_{i} J^{j}{}_{j'} \delta^i_j = \left(J^{-1}\right)^{i'}{}_{i} J^{i}{}_{j'} = \delta^{i'}_{j'} \tag{65}$$

Yes, we showed that (64) hold for any $J \in GL(n)$, so δ^i_j is a general tensor. Going back to our first definition of an abstract tensor, we say that of course, δ^i_j are the components of a $\binom{1}{1}$-tensor, namely the identity linear operator on

[11]To make this example precise we need the notion of a tensor **field**.

V or V^*. It is the linear map $id_V : V \to V$ where $id_V(v) = v$,[12] and the matrix components of the identity map are $diag(1, ..., 1)$ with respect to any basis. In fact, this is the only $\begin{pmatrix}1\\1\end{pmatrix}$-tensor whose components are invariant under any basis transformation[13]. Notice that δ_{ij} is a different kind of tensor, or different kind of map.

Exercise: Is the Levi-Civita symbol a tensor?

The Levi-Civita symbol is defined as the array

$$\varepsilon_{a_1a_2...a_n} = \begin{cases} +1 & if \ (a_1a_2...a_n) \ is \ an \ even \ permutation \ of \ (123...n) \\ -1 & if \ (a_1a_2...a_n) \ is \ an \ odd \ permutation \ of \ (123...n) \\ 0 & otherwise \end{cases} \tag{66}$$

in any basis, where $n = dimV$. To put in words, for any vector space V, the Levi-Civita symbol has $dimV$ completely antisymmetric indices. Again, the Levi-Civita symbol is defined to have the same numerical values independent of the basis, so we need to check if in addition for it being invariant if it is also a tensor.

For simplicity we first look at the case of $n = 2$, where we can calculate explicitly with matrices. In two dimensions ε_{ij} has components

$$\varepsilon_{ij} = \begin{cases} +1 & if \quad (ij) = (12) \\ -1 & if \quad (ij) = (21) \\ 0 & otherwise \end{cases} \tag{67}$$

and in matrix form

$$\begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{68}$$

Let us check whether it transforms as a $\begin{pmatrix}0\\2\end{pmatrix}$-tensor. In matrix form a $\begin{pmatrix}0\\2\end{pmatrix}$-tensor transformation is $\widetilde{T} = J^T T J$. Check for ε_{ij} and a general matrix $J = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $det\,(J) = ad - bc \neq 0$.

$$\begin{aligned} J^T \varepsilon J &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} c & d \\ -a & -b \end{pmatrix} \\ &= \begin{pmatrix} 0 & ad - bc \\ bc - ad & 0 \end{pmatrix} = det\,(J) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = det\,(J)\, \widetilde{\varepsilon} \end{aligned} \tag{69}$$

[12] And the same for the dual space.

[13] And the *Generalized Kronecker Delta Symbols* are the only $\begin{pmatrix}p\\p\end{pmatrix}$-invariant -tensors.

where the two matrices $\varepsilon = \widetilde{\varepsilon} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ are the same since the symbol is defined to be the same in all bases. Putting back the indices we write the symbol in the new basis in term of the old basis

$$\varepsilon_{i'j'} = (detJ)^{-1} J^{i}_{\ i'} J^{j}_{\ j'} \varepsilon_{ij} \tag{70}$$

The definition of Levi-Civita having the same values in any basis and having maximum completely antisymmetric indices, results in the transformation (70). This is also the general result for any n, since

$$\varepsilon_{i_1 \ldots i_n} J^{i_1}_{\ i'_1} \ldots J^{i_n}_{\ i'_n} = det(J) \varepsilon_{i'_1 \ldots i'_n} \tag{71}$$

This is the transformation of a $\begin{pmatrix} 0 \\ n \end{pmatrix}$-tensor up to a scale factor of the determinant of the transformation matrix:

$$\varepsilon_{i'_1 \ldots i'_n} = (detJ)^{-1} J^{i_1}_{\ i'_1} \ldots J^{i_n}_{\ i'_n} \varepsilon_{i_1 \ldots i_n} \tag{72}$$

Therefore - the Levi-Civita symbol is **not** a tensor. Nevertheless, it does resemble the tensor transformation and this factor of $(detJ)^{\omega}$ for some integer ω is common. An object with such transformation law is called a *tensor density of weight* ω.[14] Levi-Civita is a tensor density of weight -1. Furthermore, if we restrict J to be an orthogonal transformation then $detJ = +1$ for proper rotations and $detJ = -1$ for improper rotations (with reflection for odd dimensions). Therefore Levi-Civita is an $O(n)$-pseudotensor and a $SO(n)$-tensor.

What is the geometrical interpretation? $det\,(v_1 \ldots v_n) = \varepsilon_{i_1 \ldots i_n} v_1^{i_1} \ldots v_n^{i_n}$ is the volume of the parallelepiped formed by the n vectors $\{v_1, \ldots v_n\}$. Transforming the vectors and ε (which is invariant) result in a new volume - the old one times the determinant of the transformation matrix. This is because ε is not a tensor, so $\varepsilon_{i_1 \ldots i_n} v_1^{i_1} \ldots v_n^{i_n}$ is not a scalar (invariant). However, an orthogonal transformation does not change the volume (up to a sign), so in this case the volume is invariant and ε is a $O(n)$-(pseudo)tensor. To conclude, in order to have an invariant volume we need to make Levi-Civita into a tensor. Instead of the canonical choice for $\varepsilon_{i_1 \ldots i_n}$ to be the same in all basis, we would want it to change so it could be a tensor and make an invariant volume element. This is extra information we need to provide for the pure vector space, i.e., endow it with some extra structure.

This example demonstrate that this definition of a tensor by its transformation law has some flexibility. It can be easily restricted but also generalized to define new kinds of objects, such as tensor densities.

[14]Some use an opposite convention for the definition of the weight, with $(detJ)^{-\omega}$.

Appendix

4 The Metric Tensor

4.1 The Metric Tensor as an Inner Product

A *pseudo inner product* on a real vector space V is a map

$$(\ , \ ) : V \times V \to \mathbb{R} \tag{73}$$

that satisfies the following conditions:

1. Bilinear

$$\begin{aligned} (av + bu, w) &= a\,(v, w) + b\,(u, w) \\ (w, av + bu) &= a\,(w, v) + b\,(w, u) \end{aligned} \tag{74}$$

2. Symmetric

$$(v, u) = (u, v) \tag{75}$$

3. Non - degenerate

$$if \ \ (v, u) = 0 \ \forall v \ \Rightarrow u = 0 \tag{76}$$

For a genuine real inner product (like the usual dot product), non-degeneracy is replaced by positive-definiteness, i.e., $(v, v) > 0$ for any non-zero vector and $(v, v) = 0$ only for the zero vector. A pseudo inner product is more general, non-zero vectors can have a zero quadrance, and also negative quadrance. The feature that still remains in (76) is that the zero vector is the only vector which is orthogonal to any other vector.

We can view the inner product in three ways: As a map "$(\ , \ )$" (as defined above); As a product "$\cdot$" ("scalar product"); Or as an object by itself (like we do with functions), name it "g". What kind of object would it be? Well, a billinear map that takes two vectors and returns a scalar is a $\binom{0}{2}$-tensor. Therefore, a **choice** of a non-degenerate symmetric $\binom{0}{2}$-tensor, called a *metric tensor*, **is** an inner product on V. The Minkowski metric tensor η is the inner product of Minkowski vector space

$$v \cdot u \equiv (v, u) \equiv \eta\,(v, u) = \eta_{\mu\nu} v^{\mu} u^{\nu} = -v^0 u^0 + v^1 u^1 + v^2 u^2 + v^3 u^3 \tag{77}$$

The quadrance of a vector is

$$q\,(v) = v \cdot v = \eta\,(v, v) = \eta_{\mu\nu} v^{\mu} v^{\nu} = -\left(v^0\right)^2 + \left(v^1\right)^2 + \left(v^2\right)^2 + \left(v^3\right)^2 \tag{78}$$

In Minkowski space there are three types of vectors, classified by their quadrance

$$\begin{aligned} &\eta\,(v, v) < 0 \ \ \textit{timelike} \\ &\eta\,(v, v) = 0 \ \ \textit{null / lightlike} \\ &\eta\,(v, v) > 0 \ \ \textit{spacelike} \end{aligned} \tag{79}$$

The normalization of a timelike vector is to -1 and the normalization of a spacelike vector is to $+1$.

4.2 Raising and Lowering of Indices

The non-degeneracy of the metric tensor (76) means that the matrix of $\eta_{\mu\nu}$ has non-zero determinant, no zero elements on the diagonal in the canonical form ($\eta_{\mu\nu} = diag\,(-1,1,1,1)$), i.e., it has an inverse. The metric tensor η takes a vector and returns a covector, and the *inverse metric tensor* η^{-1} takes a covector and returns a vector. η^{-1} is a $\binom{2}{0}$-tensor, denoted as $\eta^{\mu\nu}$, with the property

$$\eta^{\mu\rho}\eta_{\rho\nu} = \delta^{\mu}_{\nu} \tag{80}$$

$\eta^{\mu\nu}$ is the induced inner product on the dual space. The metric tensor provides a unique identification between vectors and covectors ("contravariant vectors" and "covariant vectors", "upper index vectors" and "lower index vectors").

The covector *dual* to the vector v^{μ} is

$$v_{\mu} \equiv \eta_{\mu\nu}v^{\nu} \tag{81}$$

Lowering an index with the metric

The contraction (summation) with either index of $\eta_{\mu\nu}$ is the same because the metric is symmetric. We say that we "lower the index" with the metric.

The vector *dual* to the covector ω_{μ} is

$$\omega^{\mu} \equiv \eta^{\mu\nu}\omega_{\nu} \tag{82}$$

Raising an index with the inverse metric

We say that we "raise the index" with the inverse metric.

The scalar product $v \cdot u$ can be written as

$$v_{\nu}u^{\nu} = v^{\mu}\eta_{\mu\nu}u^{\nu} = v^{\mu}u_{\mu} \tag{83}$$

where on the left we lowered the index of v and on the right we lowered the index of u. In abstract terms, the covector dual to v acting on the vector u, equals the covector dual to u acting on the vector v, equals the inner product between the vectors v and u. In matrix notation (83) reads

$$\begin{pmatrix} \quad v_{\nu} \quad \end{pmatrix}\begin{pmatrix} \\ u^{\nu} \\ \\ \end{pmatrix} = \begin{pmatrix} \quad v^{\mu} \quad \end{pmatrix}\begin{pmatrix} & & \\ & \eta_{\mu\nu} & \\ & & \end{pmatrix}\begin{pmatrix} \\ u^{\nu} \\ \\ \end{pmatrix} = \begin{pmatrix} \quad v^{\mu} \quad \end{pmatrix}\begin{pmatrix} \\ u_{\mu} \\ \\ \end{pmatrix} \tag{84}$$

4.3 Lorentz Transformations Preserve the Inner Product

Consider a linear operator on V

$$\begin{aligned} \Lambda : V &\to V \\ v &\to v' = \Lambda v \end{aligned} \tag{85}$$

Λ is called an *orthogonal operator* if it preserves the inner product between any two vectors v, u (thus also preserves the quadrance of any vector)

$$\eta\left(\Lambda v, \Lambda u\right) = \eta\left(v, u\right) \tag{86}$$

In components (86) reads

$$\eta_{\mu'\nu'}\left(\Lambda v\right)^{\mu'}\left(\Lambda u\right)^{\nu'} = \eta_{\mu'\nu'}\Lambda^{\mu'}{}_{\rho}v^{\rho}\Lambda^{\nu'}{}_{\sigma}u^{\sigma} = \eta_{\rho\sigma}v^{\rho}u^{\sigma} \tag{87}$$

(87) should be satisfied for any v and u , so we can write this condition on Λ as

$$\Lambda^{\mu'}{}_{\rho}\Lambda^{\nu'}{}_{\sigma}\eta_{\mu'\nu'} = \eta_{\rho\sigma} \tag{88}$$

Preserving the inner product between any two vectors is the same as preserving the metric tensor. Such Λ is called a Lorentz transformation, and an object that under Lorentz transformation transforms as

$$v^{\mu'} = \Lambda^{\mu'}{}_{\nu}v^{\nu} \tag{89}$$

is called a 4-vector.