# Tensor Basics *

Daniel Hug

August – November 2023

## 1 Tensor Basics

*Definition 1:* A tensor is an object that is invariant under change of the coordinate system basis vectors, and has components that change in a special, predictable way under a change of coordinates, i.e. a change of the basis.

*Definition 2:* A tensor is a collection of vectors and covectors combined together using the tensor product.

*Sources: youtube channel "eigengchris"

## 1.1 Basis Transformation Rules

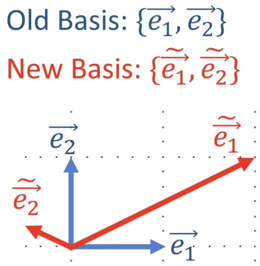



Figure 1: old and new basis vectors

*Basis transformation:* The basis vectors of the new system expressed as a linear combination of the basis vectors of the old system is considered a forward transformation. The other direction is considered a backward transformation.

Here $F$ stands for <u>F</u>orward transformation. This transformation maps the basis vectors of the old basis to the new basis vectors. The coordinates of the mapped basis vectors in the old system become the columns of the resulting forward transformation matrix. Basis vectors are chosen to be left multiplied to the transformation matrix, which corresponds to summation via the first index of the transformation matrix. Indices are chosen upper and lower to enable the Einstein summation convention later on. Transformation indices are always from “north west” to “south east”:

$$\begin{aligned}
\left[\overrightarrow{\widetilde{e_1}} \quad \overrightarrow{\widetilde{e_2}}\right] &= \left[\overrightarrow{e_1} \quad \overrightarrow{e_2}\right] \begin{bmatrix} F^1{}_1 & F^1{}_2 \\ F^2{}_1 & F^2{}_2 \end{bmatrix} \\
\overrightarrow{\widetilde{e_1}} &= F^1{}_1 \overrightarrow{e_1} + F^2{}_1 \overrightarrow{e_2} \underset{\text{figure 1}}{=} 2\overrightarrow{e_1} + 1\overrightarrow{e_2} \\
\overrightarrow{\widetilde{e_2}} &= F^1{}_2 \overrightarrow{e_1} + F^2{}_2 \overrightarrow{e_2} \underset{\text{figure 1}}{=} -\tfrac{1}{2}\overrightarrow{e_1} + \tfrac{1}{4}\overrightarrow{e_2} \\
\Rightarrow \widetilde{e_j} &= F^k{}_j \, e_k \quad \text{(summation convention)}
\end{aligned} \tag{1}$$

Here $B$ stands for <u>B</u>ackward transformation:

$$\begin{aligned}
\left[\overrightarrow{e_1} \quad \overrightarrow{e_2}\right] &= \left[\overrightarrow{\widetilde{e_1}} \quad \overrightarrow{\widetilde{e_2}}\right] \begin{bmatrix} B^1{}_1 & B^1{}_2 \\ B^2{}_1 & B^2{}_2 \end{bmatrix} \\
\overrightarrow{e_1} &= B^1{}_1 \overrightarrow{\widetilde{e_1}} + B^2{}_1 \overrightarrow{\widetilde{e_2}} \underset{\text{figure 1}}{=} \tfrac{1}{4}\overrightarrow{\widetilde{e_1}} + (-1)\overrightarrow{\widetilde{e_2}} \\
\overrightarrow{e_2} &= B^1{}_2 \overrightarrow{\widetilde{e_1}} + B^2{}_2 \overrightarrow{\widetilde{e_2}} \underset{\text{figure 1}}{=} \tfrac{1}{2}\overrightarrow{\widetilde{e_1}} + 2\overrightarrow{\widetilde{e_2}} \\
\Rightarrow e_i &= B^j{}_i \, \widetilde{e_j} \quad \text{(summation convention)}
\end{aligned} \tag{2}$$

Between $B$ and $F$ following relation holds

$$\begin{aligned} B^j{}_i F^k{}_j &= F^k{}_j B^j{}_i = \delta_i^k = \begin{cases} 1, & \text{if } i = k \\ 0, & \text{if } i \neq k \end{cases} \\ \text{full equation transposed:} &= B^i{}_j F^j{}_k = \delta_k^i \\ \Rightarrow B &= F^{-1} \end{aligned} \tag{3}$$

i.e. they are inverses to each other.

## 1.2 Vector Definition

- Vectors are invariant under a change of the coordinate system.
- Vector components are not invariant under a change of the coordinate system.
- Vectors can be defined for spaces of arbitrary dimension $n$. They form the vector space $V$ with their components $\in \mathbb{R}^n$.
- Not all vectors can be visualized as arrows (only easily done for vectors in Euclidean spaces).

Vectors are members of a vector space. A vector space needs four things:

1. $V$: a set of vectors
2. $S$: a set of scalars
3. $+$: an operation to add vectors $\vec{v} + \vec{w} = \overrightarrow{v + w}$
4. $\cdot$: an operation to scale a vector by a scalar, e.g. $2\vec{v}$

Vectors are basically things that can be added together and scaled by a factor. They are written as colums of numbers of their components in the respective basis.

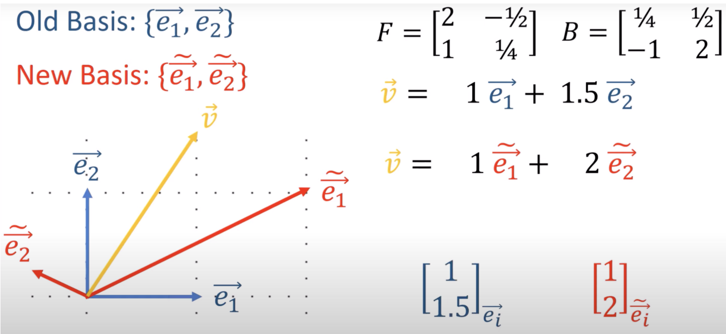



Figure 2: vector example in old and new system

Vectors are formed by linear combinations of their basis vectors with the coordinates as scaling coefficients of the corresponding basis vectors:

$$\vec{v} = \sum_i v^i \vec{e_i} = v^i e_i \quad \text{(summation convention)} \tag{4a}$$

$$\vec{v} = \sum_i \tilde{v}^i \vec{\tilde{e}_i} = \tilde{v}^i \tilde{e}_i \quad \text{(summation convention)} \tag{4b}$$

Inserting the backward transformation (2) in (4a) and re-arranging the sums by exchanging the indices, we can get the transformation rule of the vector components by comparing with (4b). This yields the transformation rule for the vector components:

$$\begin{aligned} \vec{v} &= v^i e_i = v^i B^j{}_i \tilde{e}_j = v^j B^i{}_j \tilde{e}_i \\ \vec{v} &= \tilde{v}^i \tilde{e}_i \\ \Rightarrow \tilde{v}^i &= B^i{}_j v^j \end{aligned} \tag{5}$$

By inserting the forward transformation (1) in (4b) and re-arranging the sums, we can get the transformation rule of the vector components by comparing with (4a). This yields the transformation rule for the vector components in the other direction:

$$\begin{aligned} \vec{v} &= v^i e_i \\ \vec{v} &= \tilde{v}^i \tilde{e}_i = \tilde{v}^i F^j{}_i e_j = \tilde{v}^j F^i{}_j e_i \\ \Rightarrow v^i &= F^i{}_j \tilde{v}^j \end{aligned} \tag{6}$$

Summarizing the transformation rules for basis vectors and the respective vector compontents we get:

$$\begin{aligned} \tilde{e}_i &= F^j{}_i e_j = e_j F^j{}_i & \qquad \tilde{v}^i &= B^i{}_j v^j \\ e_i &= B^j{}_i \tilde{e}_j = \tilde{e}_j B^j{}_i & \qquad v^i &= F^i{}_j \tilde{v}^j \end{aligned} \tag{7}$$

As can be seen, the basis vectors and their componentens transform with different transformations. This transformation behaviour is called contravariant. Thus vectors are contravariant tensors.

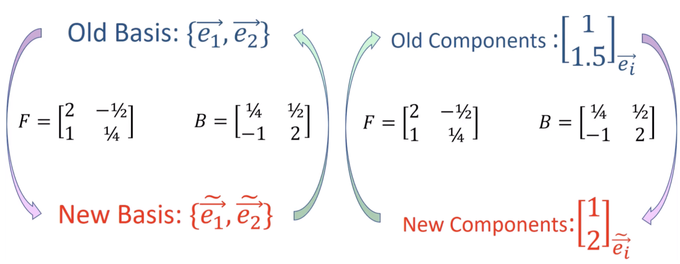



Figure 3: Summary of transformations (example)

For basis transformations the basis vectors are applied to the transformation matrix from the left whereas vector components are multiplied to the transformation matrix from the right. This can be seen in the index notation as well by the position of the indices. Multiplication from the left corresponds to summation over the first index, while multiplication from the right corresponds to summation over the second index of the transformation matrix.

## 1.3 Covector Definition:

Covectors and vectors are fundamentally different objects[1]. Covectors are functions that take vectors as arguments. They yield a scalar output $\alpha$ on a vector input $\vec{v}$:

$$\alpha(\vec{v}) = \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_n v^n = \sum_i \alpha_i v^i \tag{8}$$

Covectors take an input from vector space $V$ and return a real number $\alpha$, i.e. $\alpha : V \to \mathbb{R}$. Covectors have the linearity properties, i.e. they are linear functions:

$$\begin{aligned} \alpha(\vec{v} + \vec{w}) &= \alpha(\vec{v}) + \alpha(\vec{w}) \\ \alpha(n\vec{v}) &= n\alpha(\vec{v}) \\ \alpha(n\vec{v} + m\vec{w}) &= n\alpha(\vec{v}) + m\alpha(\vec{w}) \end{aligned} \tag{9}$$

Covectors can also be viewed as elements of their own vector space $V^*$, i.e. they can be scaled and added:

$$\begin{aligned} (n\alpha)(\vec{v}) &= n\alpha(\vec{v}) \\ (\beta + \gamma)(\vec{v}) &= \beta(\vec{v}) + \gamma(\vec{v}) \end{aligned} \tag{10}$$

The vector space of contravariant vectors is defined by $(V, S, +, \cdot)$. The vector space of covectors is defined by $(V^*, S, +, \cdot)$. The covector operations $+$ and $\cdot$ are different from the corresponding $+$ and $\cdot$ operations of vectors.
Covectors can be visualized as oriented stacks of lines of constant function value. When the coordinates are vizualized as vectors they would provide a vector oriented as the normal to the stack height lines:

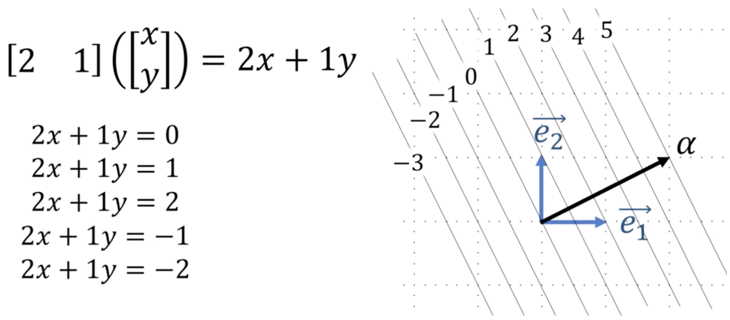



Figure 4: Example: covector visualized as lines of constant function value

[1]Covectors can "basically" be regarded as row vectors. However, only in an orthormal basis there is a trivial link to column vector components. This is not true in the general case of non-orthonormal basis vectors.

The function values returned by the covector can be visualized depending on the vectors provided as arguments. The number of height lines pierced by the vector gives the returned result:

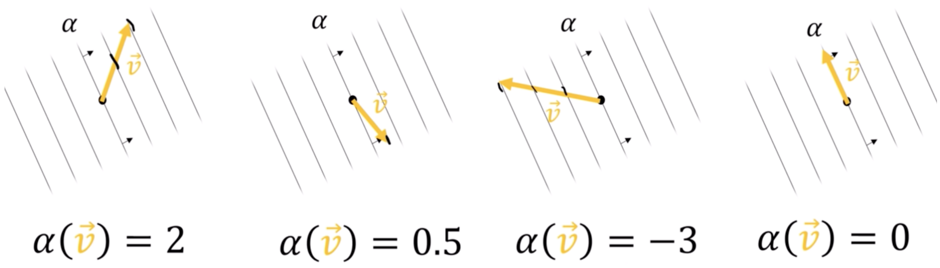



Figure 5: Example: covector return values from given input vectors

- Covectors are invariant under a change of the coordinate system.
- Covector components are not invariant under a change of the coordinate system.

Covectors are functions $\alpha : V \to \mathbb{R}$ that take vectors as arguments. Covectors don't live in the vector space $V$, but in $V^*$. Thus we cannot use basis vectors like $\{\vec{e_1}, \vec{e_2}\}$ to measure covectors directly. But still the basis $\{\vec{e_1}, \vec{e_2}\}$ can be used to derive a special covector basis $\{\epsilon^1, \epsilon^2\}$ with the specific property:

$$\begin{aligned} \epsilon^i(\vec{e_j}) &= \delta^i_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \\ \text{such that for the 2D case:} \quad \epsilon^1(\vec{e_1}) &= 1 \qquad \epsilon^1(\vec{e_2}) = 0 \\ \epsilon^2(\vec{e_1}) &= 0 \qquad \epsilon^2(\vec{e_2}) = 1 \end{aligned} \tag{11}$$

This effectively links the bases $\epsilon^i$ and $\vec{e_i}$ and the vector spaces $V$ and $V^*$, making $V^*$ the dual space of $V$. It enables us to use these covectors to find the coordinates of arbitrary vectors of $V$ by applying these special covectors to them:

$$\begin{aligned} \epsilon^1(\vec{v}) &= \epsilon^1(v^1\vec{e_1} + v^2\vec{e_2}) \underset{(9)}{=} v^1\epsilon^1(\vec{e_1}) + v^2\epsilon^1(\vec{e_2}) \underset{(11)}{=} v^1 \\ \epsilon^2(\vec{v}) &= \epsilon^2(v^1\vec{e_1} + v^2\vec{e_2}) \underset{(9)}{=} v^1\epsilon^2(\vec{e_1}) + v^2\epsilon^2(\vec{e_2}) \underset{(11)}{=} v^2 \\ \Rightarrow \epsilon^i(\vec{v}) &= v^i \end{aligned} \tag{12}$$

This means we select a special covector basis that is aligned with our vector basis such that each covector basis selects it's corresponding basis vector contribution exclusively:


Figure 6: Example: specifically chosen covector basis

If we now want to apply a general covector $\alpha$ to a vector $\vec{v}$ we can transform the equations by defining the covector components in terms of our specially chosen covector basis vectors $\epsilon^1$ and $\epsilon^2$:

$$\begin{aligned}
\alpha(\vec{v}) &= \alpha(v^1\overrightarrow{e_1} + v^2\overrightarrow{e_2}) \underset{(9)}{=} v^1\alpha(\overrightarrow{e_1}) + v^2\alpha(\overrightarrow{e_2}) \\
&\underset{(12)}{=} \epsilon^1(\vec{v})\alpha(\overrightarrow{e_1}) + \epsilon^2(\vec{v})\alpha(\overrightarrow{e_2}) \\
\text{defining}: \alpha(\overrightarrow{e_1}) &= \alpha_1 \\
\alpha(\overrightarrow{e_2}) &= \alpha_2 \\
\alpha(\vec{v}) &= \alpha_1\epsilon^1(\vec{v}) + \alpha_2\epsilon^2(\vec{v}) \\
&\underset{(10)}{=} (\alpha_1\epsilon^1 + \alpha_2\epsilon^2)(\vec{v}) \\
\Rightarrow \alpha &= \alpha_1\epsilon^1 + \alpha_2\epsilon^2 = \alpha_i\epsilon^i
\end{aligned} \tag{13}$$

Thus every covector can be expressed as a linear combination of our specifically chosen covector basis. We call the basis $\{\epsilon^1, \epsilon^2\}$ the "dual basis" of the vector space $V^*$. Fig. 7 shows this graphically.

Of course we could also define an arbitray covector with respect to the vector basis in the new system: The basis $\{\widetilde{\overrightarrow{e_1}}, \widetilde{\overrightarrow{e_2}}\}$ can be used to derive a special

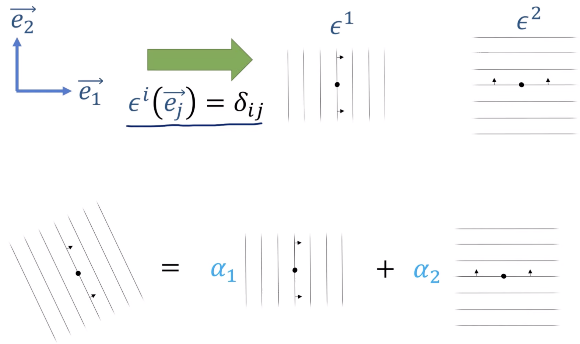



Figure 7: Arbitrary covector (lower row left) expressed as linear combination of the covector basis $\{\epsilon^1, \epsilon^2\}$ after the dual basis has been derived from the vector basis.

covector basis $\{\tilde{\epsilon}^1, \tilde{\epsilon}^2\}$ with the specific property:

$$
\begin{aligned}
\tilde{\epsilon}^i(\tilde{\vec{e_j}}) &= \delta^i_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \\
\text{such that for the 2D case:} \quad \tilde{\epsilon}^1(\tilde{\vec{e_1}}) &= 1 \qquad \tilde{\epsilon}^1(\tilde{\vec{e_2}}) = 0 \\
\tilde{\epsilon}^2(\tilde{\vec{e_1}}) &= 0 \qquad \tilde{\epsilon}^2(\tilde{\vec{e_2}}) = 1
\end{aligned}
\tag{14}
$$

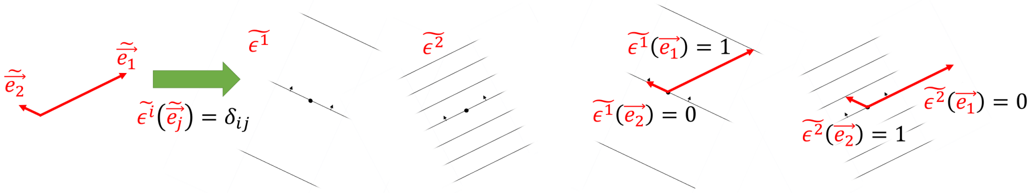
This effectively links the bases $\tilde{\epsilon}^i$ and $\tilde{\vec{e_i}}$ and the vector spaces $V$ and $V^*$, making $V^*$ the dual space of $V$. It enables us to use these covectors to find the coordinates of arbitrary vectors of $V$ by applying these special covectors to them:

$$
\begin{aligned}
\tilde{\epsilon}^1(\vec{v}) &= \tilde{\epsilon}^1(\tilde{v}^1\tilde{\vec{e_1}} + \tilde{v}^2\tilde{\vec{e_2}}) \underset{(9)}{=} \tilde{v}^1\tilde{\epsilon}^1(\tilde{\vec{e_1}}) + \tilde{v}^2\tilde{\epsilon}^1(\tilde{\vec{e_2}}) \underset{(14)}{=} \tilde{v}^1 \\
\tilde{\epsilon}^2(\vec{v}) &= \tilde{\epsilon}^2(\tilde{v}^1\tilde{\vec{e_1}} + \tilde{v}^2\tilde{\vec{e_2}}) \underset{(9)}{=} \tilde{v}^1\tilde{\epsilon}^2(\tilde{\vec{e_1}}) + \tilde{v}^2\tilde{\epsilon}^2(\tilde{\vec{e_2}}) \underset{(14)}{=} \tilde{v}^2 \\
\Rightarrow \tilde{\epsilon}^i(\vec{v}) &= \tilde{v}^i
\end{aligned}
\tag{15}
$$

This leads to another dual basis specific for the basis vectors in the new system as shown in equation 16.

$$
\begin{aligned}
\alpha(\vec{v}) &= \alpha(\widetilde{v}^1\widetilde{\vec{e_1}} + \widetilde{v}^2\widetilde{\vec{e_2}}) \underset{(9)}{=} \widetilde{v}^1\alpha(\widetilde{\vec{e_1}}) + \widetilde{v}^2\alpha(\widetilde{\vec{e_2}}) \\
&\underset{(15)}{=} \widetilde{\epsilon^1}(\vec{v})\alpha(\widetilde{\vec{e_1}}) + \widetilde{\epsilon^2}(\vec{v})\alpha(\widetilde{\vec{e_2}}) \\
\text{defining: } \alpha(\widetilde{\vec{e_1}}) &= \widetilde{\alpha_1} \\
\alpha(\widetilde{\vec{e_2}}) &= \widetilde{\alpha_2} \qquad (16) \\
\alpha(\vec{v}) &= \widetilde{\alpha_1}\widetilde{\epsilon^1}(\vec{v}) + \widetilde{\alpha_2}\widetilde{\epsilon^2}(\vec{v}) \\
&\underset{(10)}{=} (\widetilde{\alpha_1}\widetilde{\epsilon^1} + \widetilde{\alpha_2}\widetilde{\epsilon^2})(\vec{v}) \\
\Rightarrow \alpha &= \widetilde{\alpha_1}\widetilde{\epsilon^1} + \widetilde{\alpha_2}\widetilde{\epsilon^2} = \widetilde{\alpha}_i\widetilde{\epsilon^i}
\end{aligned}
$$

Graphically fig. 8 shows the covectors with respect to the basis in the new system.




(a) Covector hight lines (new basis) (b) Covector function (new basis)

Figure 8: Another dual covector basis expressed with respect to the basis in the new system

In order to derive the transformation rules for the the covector basis we use $\{\epsilon^j\}$ as the old basis for $V^*$ and $\{\widetilde{\epsilon^i}\}$ as the new basis for $V^*$. The new basis can be expressed by a linear combination of the old basis vectors using the coefficients $Q^i{}_j$:

$$\widetilde{\epsilon^i} = Q^i{}_j\epsilon^j \qquad (17)$$

Using (14) and (17) we can write:

$$\begin{aligned}
\widetilde{\epsilon^i}(\overrightarrow{\widetilde{e_k}}) &= Q^i{}_j \epsilon^j(\overrightarrow{\widetilde{e_k}}) \\
\delta^i_k &\underset{(14),(1)}{=} Q^i{}_j \epsilon^j(F^l{}_k \overrightarrow{e_l}) \underset{(9)}{=} Q^i{}_j F^l{}_k \epsilon^j(\overrightarrow{e_l}) \\
\delta^i_k &\underset{(11)}{=} Q^i{}_j F^l{}_k \delta^j_l = Q^i{}_j F^j{}_k \\
\text{comparing with (3): } Q^i{}_j &= B^i{}_j \\
\text{inserting in (17)} \Rightarrow \widetilde{\epsilon^i} &= B^i{}_j \epsilon^j
\end{aligned} \tag{18}$$

This shows that covectors transform with the backward transformation from the old to the new basis. Using the same approach it can be shown that:

$$\Rightarrow \epsilon^i = F^i{}_j \widetilde{\epsilon^j} \tag{19}$$

Summarizing the transformation rules for vectors and covectors we get:

$$\begin{aligned}
\widetilde{e_j} &= F^i{}_j e_i & \widetilde{\epsilon^i} &= B^i{}_j \epsilon^j \\
e_j &= B^i{}_j \widetilde{e_i} & \epsilon^i &= F^i{}_j \widetilde{\epsilon^j}
\end{aligned} \tag{20}$$

The covector can be constructed from the basis in the old and new basis:

$$\alpha = \alpha_i \epsilon^i = \widetilde{\alpha}_j \widetilde{\epsilon^j} \tag{21}$$

By inserting the transformation rule (18) or (19) and comparing with the other part of the equation (21) we can get the transformation rules for the covector components:

$$\begin{aligned}
\widetilde{\alpha}_j &= F^i{}_j \alpha_i \\
\alpha_j &= B^i{}_j \widetilde{\alpha}_i
\end{aligned} \tag{22}$$

This means that covector components transform as the basis vectors do, i.e. with the forward transformation from the old to the new system and with the backward transformation from the new to the old system. This is called covariant transformation behaviour. For both cases the summation goes via the first index of the transformation matrix (equivalent to matrix multiplication from the left).

Summing up everything so far we get:

$$\boxed{\begin{array}{rrcl rcl l}
 & \widetilde{e_j} &=& F^i{}_j e_i & \widetilde{\epsilon^i} &=& B^i{}_j \epsilon^j & (\underline{\text{contra}}\text{var.}) \\
 & e_j &=& B^i{}_j \widetilde{e_i} & \epsilon^i &=& F^i{}_j \widetilde{\epsilon^j} & \\
\text{with:} & \vec{v} &=& v^i e_i = \widetilde{v^j} \widetilde{e_j} & \alpha &=& \alpha_i \epsilon^i = \widetilde{\alpha}_j \widetilde{\epsilon^j} & \\
(\underline{\text{contra}}\text{var.}) & \widetilde{v^i} &=& B^i{}_j v^j & \widetilde{\alpha}_j &=& F^i{}_j \alpha_i & (\underline{\text{co}}\text{var.}) \\
 & v^i &=& F^i{}_j \widetilde{v^j} & \alpha_j &=& B^i{}_j \widetilde{\alpha}_i &
\end{array}} \tag{23}$$

## 1.4 Linear Maps

Linear maps take a vector as an input and produce a new vector as an output. Linear maps don't transform the basis vectors. They just transform the input vectors into the output vectors given in the same basis.

Linear maps shown as matrices map copies of the unit basis vectors $[1,0]^T$ and $[0,1]^T$ into the vectors their coordinate columns consist of (see matrix multiplication with the corresponding vector from the right side). The first column is the mapped value of the first basis vector and the second column the mapped value of the second basis vector. In general the $i^{th}$ column is the mapped value of the $i^{th}$ basis vector.

From a geometrical point of view linear maps are spacial transforms that keep gridlines parallel, keep gridlines evenly spaced, and keep the origin stationary. That is linear maps can stretch and rotate the input vectors, but they don't shift the origin.

From an abstract point of view linear maps do map vectors to vectors and have the linearity properties, i.e. can be added and scaled. That means:

$$\begin{array}{rcl} L : V \to V & \text{or} & L : V \to W \\ L(\vec{v} + \vec{w}) & = & L(\vec{v}) + L(\vec{w}) \\ L(n\vec{v}) & = & nL(\vec{v}) \end{array} \tag{24}$$

We can derive the linear map and the resulting matrix multiplication of the vector components directly from the linearity properties (24):

$$\begin{array}{rcl} \vec{w} = L(\vec{v}) & = & L(v^1\vec{e_1} + v^2\vec{e_2}) \\ & \underset{(24)}{=} & v^1 L(\vec{e_1}) + v^2 L(\vec{e_2}) \end{array} \tag{25}$$

Assuming $L : V \to V$, we can express the resulting vector in $\{\vec{e_1}, \vec{e_2}\}$:

$$\begin{array}{rcl} L(\vec{e_1}) & = & L^1{}_1\vec{e_1} + L^2{}_1\vec{e_2} \\ L(\vec{e_2}) & = & L^1{}_2\vec{e_1} + L^2{}_2\vec{e_2} \\ L(\vec{e_j}) & = & L^k{}_j\vec{e_k} \end{array} \tag{26}$$

$$\begin{array}{rcl} \Rightarrow \vec{w} = L(\vec{v}) & = & v^1(L^1{}_1\vec{e_1} + L^2{}_1\vec{e_2}) + v^2(L^1{}_2\vec{e_1} + L^2{}_2\vec{e_2}) \\ & = & (L^1{}_1 v^1 + L^1{}_2 v^2)\vec{e_1} + (L^2{}_1 v^1 + L^2{}_2 v^2)\vec{e_2} \\ & = & w^1 \vec{e_1} + w^2 \vec{e_2} \\ \\ w^1 & = & L^1{}_1 v^1 + L^1{}_2 v^2 \\ w^2 & = & L^2{}_1 v^1 + L^2{}_2 v^2 \\ \\ \Rightarrow w^i & = & L^i{}_j v^j \end{array} \tag{27}$$

To figure out the transformation rules for linear maps we need to express the linear map as an expansion in the new basis to see how the new basis vector are mapped:

$$
\begin{aligned}
\widetilde{L^l}_i \widetilde{\vec{e_l}} &= L(\widetilde{\vec{e_i}}) \\
&\underset{(1)}{=} L(F^j_i \vec{e_j}) \underset{(24)}{=} F^j_i L(\vec{e_j}) \\
&\underset{(26)}{=} F^j_i L^k_j \vec{e_k} \underset{(2)}{=} F^j_i L^k_j B^l_k \widetilde{\vec{e_l}} \\
\widetilde{L^l}_i \widetilde{\vec{e_l}} &= B^l_k L^k_j F^j_i \widetilde{\vec{e_l}} \\
\Rightarrow \widetilde{L^l}_i &= B^l_k L^k_j F^j_i
\end{aligned} \tag{28}
$$

To figure out the transformation to the old basis we can use the fact that the forward and the backward transform are inverses of each other (3). By multiplying (28) from the left with the forward transformation matrix $F^s_l$ and from the right with the backward transformation matrix $B^i_t$we get:

$$
\begin{aligned}
\widetilde{L^l}_i &= B^l_k L^k_j F^j_i \\
F^s_l \widetilde{L^l}_i B^i_t &= F^s_l B^l_k L^k_j F^j_i B^i_t \\
&= \delta^s_k L^k_j \delta^j_t = L^s_t \\
\Rightarrow L^s_t &= F^s_l \widetilde{L^l}_i B^i_t
\end{aligned} \tag{29}
$$

Linear transformations can also be interpreted as linear combinations of vector-covector pairs, i.e. as sum of tensor products of covariant basis vectors $\vec{e_i}$ and contravariant covector basis vectors $\epsilon^j$. They form a basis for any arbitrary linear transformation matrix, i.e. the set $\{\vec{e_1}\epsilon^1, \vec{e_1}\epsilon^2, \vec{e_2}\epsilon^1, \vec{e_2}\epsilon^2\}$ is a basis for $V \to V$ (the same consideration is possible for the new basis).

$$
\begin{aligned}
L &= a\vec{e_1}\epsilon^1 + b\vec{e_1}\epsilon^2 + c\vec{e_2}\epsilon^1 + d\vec{e_2}\epsilon^2 \\
L &= L^i_j \vec{e_i}\epsilon^j = L^i_j(\vec{e_i} \otimes \epsilon^j)
\end{aligned} \tag{30}
$$

In matrix notation the matrix for an arbitrary linear transformation is:

$$
a\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + b\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + c\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} + d\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} = a\begin{bmatrix}a & b\\c & d\end{bmatrix} \tag{31}
$$

Using the second equation of (30) and the vector definition $\vec{v} = v^k e_k$ we get:

$$
\begin{aligned}
\vec{w} &= L(\vec{v}) \\
&= L^i_j \vec{e_i}\epsilon^j(v^k e_k) \\
&= L^i_j v^k \vec{e_i}\epsilon^j(e_k) \\
&= L^i_j v^k \vec{e_i}\delta^j_k \\
\Rightarrow \vec{w} &= L^i_j v^j \vec{e_i} \\
w^i &= L^i_j v^j
\end{aligned} \tag{32}
$$

## 1.5 Metric Tensor

The length of a vector in an arbitrary coordinate system is defined by its norm which in turn is defined by the scalar product:

$$\begin{aligned}
\|\vec{v}\|^2 &= \vec{v}\cdot\vec{v} \\
&= (v^1\overrightarrow{e_1} + v^2\overrightarrow{e_2})\cdot(v^1\overrightarrow{e_1} + v^2\overrightarrow{e_2}) \\
&= (v^1)^2(\overrightarrow{e_1}\cdot\overrightarrow{e_1}) + v^1v^2(\overrightarrow{e_1}\cdot\overrightarrow{e_2}) + v^2v^1(\overrightarrow{e_2}\cdot\overrightarrow{e_1}) + (v^2)^2(\overrightarrow{e_2}\cdot\overrightarrow{e_2}) \\
&= (v^1)^2(\overrightarrow{e_1}\cdot\overrightarrow{e_1}) + 2v^1v^2(\overrightarrow{e_1}\cdot\overrightarrow{e_2}) + (v^2)^2(\overrightarrow{e_2}\cdot\overrightarrow{e_2}) \\
&= (\tilde{v}^1)^2(\widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_1}}) + 2\tilde{v}^1\tilde{v}^2(\widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_2}}) + (\tilde{v}^2)^2(\widetilde{\overrightarrow{e_2}}\cdot\widetilde{\overrightarrow{e_2}})
\end{aligned} \tag{33}$$

In orthogonal coordinate systems the mixed products $(\overrightarrow{e_1}\cdot\overrightarrow{e_2})$ and $(\widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_2}})$ are equal to 0, but in the more general non-orthogonal case these terms do not disappear.

The scalar product can be written in component form based on the summation convention with an implied summation over $i$ and $j$ as

$$\begin{aligned}
\|\vec{v}\|^2 = \vec{v}\cdot\vec{v} &= (v^i\overrightarrow{e_i})\cdot(v^j\overrightarrow{e_j}) = v^iv^j(\overrightarrow{e_i}\cdot\overrightarrow{e_j}) = v^iv^jg_{ij} \\
&= (\tilde{v}^i\widetilde{\overrightarrow{e_i}})\cdot(\tilde{v}^j\widetilde{\overrightarrow{e_j}}) = \tilde{v}^i\tilde{v}^j(\widetilde{\overrightarrow{e_i}}\cdot\widetilde{\overrightarrow{e_j}}) = \tilde{v}^i\tilde{v}^j\tilde{g}_{ij}
\end{aligned} \tag{34}$$

Here $g$ is the metric tensor given in coordinates of the old and new system respectively. It's components are defined by the scalar products:

$$\begin{aligned}
g_{ij} &= (\overrightarrow{e_i}\cdot\overrightarrow{e_j}) &= (\overrightarrow{e_j}\cdot\overrightarrow{e_i}) &= g_{ji} \\
\tilde{g}_{ij} &= (\widetilde{\overrightarrow{e_i}}\cdot\widetilde{\overrightarrow{e_j}}) &= (\widetilde{\overrightarrow{e_j}}\cdot\widetilde{\overrightarrow{e_i}}) &= \tilde{g}_{ji}
\end{aligned} \tag{35}$$

In component form it can be written as a symmetric matrix, because the sequence of arguments does not influence the result of the dot product.

$$\begin{aligned}
g_{\overrightarrow{e_i}} &= \begin{bmatrix} \overrightarrow{e_1}\cdot\overrightarrow{e_1} & \overrightarrow{e_1}\cdot\overrightarrow{e_2} \\ \overrightarrow{e_2}\cdot\overrightarrow{e_1} & \overrightarrow{e_2}\cdot\overrightarrow{e_2} \end{bmatrix}_{\overrightarrow{e_i}} = \begin{bmatrix} \overrightarrow{e_1}\cdot\overrightarrow{e_1} & \overrightarrow{e_1}\cdot\overrightarrow{e_2} \\ \overrightarrow{e_1}\cdot\overrightarrow{e_2} & \overrightarrow{e_2}\cdot\overrightarrow{e_2} \end{bmatrix}_{\overrightarrow{e_i}} \\
g_{\widetilde{\overrightarrow{e_i}}} &= \begin{bmatrix} \widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_1}} & \widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_2}} \\ \widetilde{\overrightarrow{e_2}}\cdot\widetilde{\overrightarrow{e_1}} & \widetilde{\overrightarrow{e_2}}\cdot\widetilde{\overrightarrow{e_2}} \end{bmatrix}_{\widetilde{\overrightarrow{e_i}}} = \begin{bmatrix} \widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_1}} & \widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_2}} \\ \widetilde{\overrightarrow{e_1}}\cdot\widetilde{\overrightarrow{e_2}} & \widetilde{\overrightarrow{e_2}}\cdot\widetilde{\overrightarrow{e_2}} \end{bmatrix}_{\widetilde{\overrightarrow{e_i}}}
\end{aligned} \tag{36}$$

For the special case of an orthonormal system these matrices become the identity matrix. All dot products can be computed in an arbitrary basis system by using the metric tensor like this:

$$\begin{aligned}
\vec{v}\cdot\vec{v} &= \|\vec{v}\|^2 &= v^iv^jg_{ij} = \tilde{v}^i\tilde{v}^j\tilde{g}_{ij} \\
\vec{w}\cdot\vec{w} &= \|\vec{w}\|^2 &= w^iw^jg_{ij} = \tilde{w}^i\tilde{w}^j\tilde{g}_{ij} \\
\vec{v}\cdot\vec{w} &= \|\vec{v}\|\|\vec{w}\|\cos(\theta) &= v^iw^jg_{ij} = \tilde{v}^i\tilde{w}^j\tilde{g}_{ij}
\end{aligned} \tag{37}$$

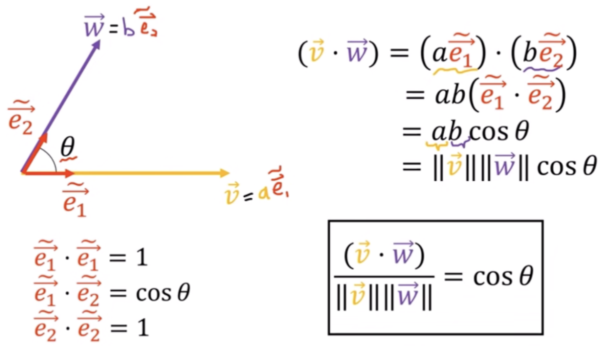



Figure 9: Angle $\theta$ between two arbitrary vectors.

To show that the last formula in equation (37) is correct, consider figure 9.

To check how the metric tensor transforms, we use the formulation of the metric tensor in the new system and transform it by replacing the tilde-components by expressing them in the old system via the definitions of the forward transform given by equation (7):

$$\begin{aligned}
\tilde{g}_{ij} &= \tilde{\vec{e}_i}\cdot\tilde{\vec{e}_j} \\
&= (F^k_{\ i}\,\vec{e_k})\cdot(F^l_{\ j}\,\vec{e_l}) \\
&= F^k_{\ i}\,F^l_{\ j}(\vec{e_k}\cdot\vec{e_l}) \\
\Rightarrow \tilde{g}_{ij} &= F^k_{\ i}\,F^l_{\ j}\,g_{kl} \\
g_{kl} &= \vec{e_k}\cdot\vec{e_l} \\
&= (B^i_{\ k}\,\tilde{\vec{e_i}})\cdot(B^j_{\ l}\,\tilde{\vec{e_j}}) \\
&= B^i_{\ k}B^j_{\ l}\,(\tilde{\vec{e_i}}\cdot\tilde{\vec{e_j}}) \\
\Rightarrow g_{kl} &= B^i_{\ k}B^j_{\ l}\,\tilde{g}_{ij}
\end{aligned} \tag{38}$$

Metric tensors are (0,2)-tensors. They transform with covariant transformation behaviour using two consecutive transformations for a change of coordinates.

The metric tensor is a function that takes two vectors as input and transforms them into a real number, thus it is an example of a so-called bilinear form: $g\colon V\times V \to \mathbb{R}$, i.e. $g(\vec{v},\vec{w}) \to v^i w^j g_{ij}$. When we put in the same vector twice we get the vector length. When we put in two vectors we get the product of their vector lengths times the cosine of the angle between them (see eq. (37)).

The metric tensor $g$ is called a bilinear form, because it takes two vectors as input, maps them to a real number as output, and has the linearity properties for both of its arguments:

$$
\begin{aligned}
g\colon V \times V \to \mathbb{R}, &\qquad g(\vec{v}, \vec{w}) \to v^i w^j g_{ij} \\
a(v^i w^j g_{ij}) &= (av^i) w^j g_{ij} = v^i (aw^j) g_{ij} \\
\Rightarrow ag(\vec{v}, \vec{w}) &= g(a\vec{v}, \vec{w}) = g(\vec{v}, a\vec{w}) \\
(v^i + u^i) w^j g_{ij} &= v^i w^j g_{ij} + u^i w^j g_{ij} \\
\Rightarrow g(\vec{v} + \vec{u}, \vec{w}) &= g(\vec{v}, \vec{w}) + g(\vec{u}, \vec{w}) \\
v^i (w^j + t^j) g_{ij} &= v^i w^j g_{ij} + v^i t^j g_{ij} \\
\Rightarrow g(\vec{v}, \vec{w} + \vec{t}) &= g(\vec{v}, \vec{w}) + g(\vec{v}, \vec{t})
\end{aligned}
\tag{39}
$$

In general forms are functions that take an arbitrary number of vector inputs and provide a scalar as an output, i.e. forms: $V \times V \times \cdots \times V \to \mathbb{R}$. Forms must have the linearity properties. Examples of a 1-form or a linear form are covectors ($V \to \mathbb{R}$). An example of a 2-form or bilinear form is the metric tensor, as shown above ($V \times V \to \mathbb{R}$).

The metric tensor is a special bilinear form, because it has two additional properties that bilinear forms do not have in general:

$$
\begin{aligned}
g(\vec{v}, \vec{w}) &= v^i w^j g_{ij} = v^i w^j g_{ji} = g(\vec{w}, \vec{v}) \quad \text{(symmetry)} \\
g(\vec{v}, \vec{v}) &= \|v\|^2 \geq 0 \quad \text{(positive vector length)}
\end{aligned}
\tag{40}
$$

Bilinear forms (including the metric tensor) can alternatively be interpreted as linear combinations of covector-covector pairs $\mathcal{B} = \mathcal{B}_{ij} \epsilon^i \epsilon^j = \mathcal{B}_{ij} (\epsilon^i \otimes \epsilon^j)$. As for linear maps, we can show that the bilinear map can be created from that definition and linearity:

$$
\begin{aligned}
\mathcal{B} &= \mathcal{B}_{kl} \epsilon^k \epsilon^l \\
&= \mathcal{B}_{kl} (F^k{}_i \widetilde{\epsilon^i})(F^l{}_j \widetilde{\epsilon^j}) \\
&= (F^k{}_i F^l{}_j \mathcal{B}_{kl}) \widetilde{\epsilon^i} \widetilde{\epsilon^j} \\
&= \widetilde{\mathcal{B}_{ij}} \widetilde{\epsilon^i} \widetilde{\epsilon^j} \\
\Rightarrow \widetilde{\mathcal{B}_{ij}} &= F^k{}_i F^l{}_j \mathcal{B}_{kl} \\
\text{and correspondingly:} \qquad \mathcal{B}_{ij} &= B^k{}_i B^l{}_j \widetilde{\mathcal{B}_{kl}}
\end{aligned}
\tag{41}
$$

Their coordinate forms can be derived from this definition $\mathcal{B} = \mathcal{B}_{ij}\epsilon^i\epsilon^j$ and the vector definitions $\vec{v} = v^k\overrightarrow{e_k}$ and $\vec{w} = w^l\overrightarrow{e_l}$ by inserting them:

$$\begin{aligned} s &= \mathcal{B}(\vec{v}, \vec{w}) \\ &= \mathcal{B}_{ij}\epsilon^i\epsilon^j(v^k\overrightarrow{e_k}, w^l\overrightarrow{e_l}) \\ &= \mathcal{B}_{ij}\epsilon^i(v^k\overrightarrow{e_k})\epsilon^j(w^l\overrightarrow{e_l}) \\ &= \mathcal{B}_{ij}v^k w^l\epsilon^i(\overrightarrow{e_k})\epsilon^j(\overrightarrow{e_l}) \\ &= \mathcal{B}_{ij}v^k w^l\delta^i_k\delta^j_l \\ &= \mathcal{B}_{ij}v^i w^j \end{aligned} \tag{42}$$

The metric tensor can also be used to describe a correspondence of vectors in the vector space $V$ to the dual space $V^*$. The correspondence exists between vectors and the scalar product of $\vec{v}$ with an open slot to be filled with another vector. This leads to a formula for lowering the index of a vector component with the help of the metric tensor $g_{ij}$ and a formular for raising the index with the help of the inverse metric tensor $\mathfrak{g}^{ij}$:

$$\begin{array}{rrcl rcl} & v_i &=& g_{ij}v^j & \widetilde{v_i} &=& \widetilde{g_{ij}}\widetilde{v^j} \\ & v^i &=& \mathfrak{g}^{ij}v_j & \widetilde{v^i} &=& \widetilde{\mathfrak{g}^{ij}}\widetilde{v_j} \\ \text{where} & \mathfrak{g}^{ki}g_{ij} &=& \delta^k_j & \widetilde{\mathfrak{g}^{ki}}\widetilde{g_{ij}} &=& \delta^k_j \end{array} \tag{43}$$

## 1.6 Transformation laws of arbitrary tensors

Let's review the defintion 1 of a tensor from the front page: *"A tensor is an object that is invariant under a change of coordinates, and has components that change in a special, predictable way under a change of coordinates."*

**To sum up everything up to here:**

Contravariant (1,0)-tensors transform like this (basis / components) - the components can be written as column vectors and are multiplied to the matrix from the right (i.e. summation over the second index):

$$\begin{aligned} \widetilde{\epsilon^i} &= B^i{}_j \epsilon^j \qquad & \widetilde{v^i} &= B^i{}_j v^j \\ \epsilon^i &= F^i{}_j \widetilde{\epsilon^j} \qquad & v^i &= F^i{}_j \widetilde{v^j} \end{aligned} \tag{44}$$

Covariant (0,1)-tensors transform like this (basis / components) - the components can be written as row vectors and are multiplied to the matrix from the left (i.e. summation over the first index):

$$\begin{aligned} \widetilde{\vec{e_j}} &= \vec{e_i} F^i{}_j = F^i{}_j \vec{e_i} \qquad & \widetilde{\alpha_j} &= \alpha_i F^i{}_j = F^i{}_j \alpha_i \\ \vec{e_j} &= \widetilde{\vec{e_i}} B^i{}_j = B^i{}_j \widetilde{\vec{e_i}} \qquad & \alpha_j &= \widetilde{\alpha_i} B^i{}_j = B^i{}_j \widetilde{\alpha_i} \end{aligned} \tag{45}$$

Linear maps are (1,1)-tensors, i.e. they have a vector and a covector part and transform like this - the matrix of the coefficients of the linear map is multiplied from the left and the right with the respective forward or backward transformation matrices:

$$\begin{aligned} \widetilde{L^i{}_j} &= B^i{}_k L^k{}_l F^l{}_j \\ L^i{}_j &= F^i{}_k \widetilde{L^k{}_l} B^l{}_j \end{aligned} \tag{46}$$

Metric tensors are (0,2)-tensors, i.e. they have two covector parts and transform like this:

$$\begin{aligned} \tilde{g}_{ij} &= F^k_i F^l{}_j g_{kl} \\ g_{kl} &= B^i{}_k B^j{}_l \tilde{g}_{ij} \end{aligned} \tag{47}$$

**The general transformation laws for tensor components for an arbitray tensor of type $(m, n)$ with $m$ upstairs indices and $n$ downstairs indices are:**

$$\begin{aligned} \widetilde{T}^{abc\ldots}_{xyz\ldots} &= (B^a{}_i B^b{}_j B^c{}_k \cdots) T^{ijk\ldots}_{rst\ldots} (F^r{}_x F^s{}_y F^t{}_z \cdots) \\ T^{ijk\ldots}_{rst\ldots} &= (F^i{}_a F^j{}_b F^k{}_c \cdots) \widetilde{T}^{abc\ldots}_{xyz\ldots} (B^x{}_r B^y{}_s B^z{}_t \cdots) \end{aligned} \tag{48}$$

The upstairs indices represent the contravariant components of the tensor and the downstairs indices represent the covariant components of the tensor. The transformation itself is a series of backward and forward transforms. To go from old to new system the covariant componens use forward transformations, while the contravariant components use backward transformations. For the transformation from the new to the old system the situation reverses.

**Anything that follows these transformation rules during change of coordinates is a tensor.**